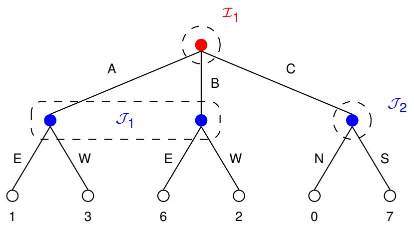
Problem 1. Behavioral strategies (4 points)

We aim to compute the behavioral saddle-point equilibrium of the zero-sum extensive form game shown below.

a) Verify that $\left(\begin{bmatrix}\frac{2}{3}\\ \frac{1}{3}\end{bmatrix}, \begin{bmatrix}\frac{1}{6}\\ \frac{5}{6}\end{bmatrix}\right)$ is a mixed-strategy equilibrium of $\begin{bmatrix}1 & 3\\ 6 & 2\end{bmatrix}$. What is the value of the game? (2 points)

b) Determine optimal strategy of player 2 (maximizer) at $\mathcal{J}_2$. (0.5 point)

c) Using the above two steps, determine the mixed strategy of player 1 at $\mathcal{I}_1$ and the strategy of player 2 for each of his information sets to characterize the behavioral saddle-point equilibrium of the game (1.5 points).



Solution:

a) we can check whether any player has any incentive in deviating. The rewards for player 2 are

$$\begin{bmatrix}\frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1 & 3\\ 6 & 2\end{bmatrix} = \begin{bmatrix}\frac{2+6}{3} & \frac{6+2}{3}\end{bmatrix} = \begin{bmatrix}\frac{8}{3}\\ \frac{8}{3}\end{bmatrix}$$

No matter which policies the second player chooses, the reward is $\frac{8}{3}$. For player 1 we have

$$\begin{bmatrix}1 & 3\\ 6 & 2\end{bmatrix}\begin{bmatrix}\frac{1}{6}\\ \frac{5}{6}\end{bmatrix} = \begin{bmatrix}\frac{1+15}{6} & \frac{6+10}{6}\end{bmatrix} = \begin{bmatrix}\frac{8}{3} & \frac{8}{3}\end{bmatrix}$$

Also the first player has no incentive in chaning the policy. The value of the game is $\frac{8}{3}$.

b) The optimal strategy of player 2 (maximizer) at $\mathcal{J}_2$ is to always play S, thus $\begin{bmatrix}0\\ 1\end{bmatrix}$.

c) Player 1 has no incentive in playing C, as the reward would be 7, bigger than any reward that could obtain by playing A or B. The probability with which should play A or B are the one provided in question a, as we proved. Thus, its mixed strategy is: $\begin{bmatrix}\frac{2}{3} & \frac{1}{3} & 1\end{bmatrix}^\top$. Player 2 should use the mixed strategy $\begin{bmatrix}\frac{1}{6} & \frac{5}{6}\end{bmatrix}^\top$ in $\mathcal{J}_1$ and $\begin{bmatrix}0 & 1\end{bmatrix}^\top$ in $\mathcal{J}_2$.

Problem 2. A zero-sum LQ game (7 points)

Consider a zero-sum linear quadratic game with two agents over two time steps. The state dynamics is

$$x_{t+1} = f(x_t, u_t, w_t) = \frac{1}{2}x_t + u_t + w_t, \ \forall t = 0, 1,$$

where $u_t \in \mathbb{R}$ and $w_t \in \mathbb{R}$ denote the control actions of the minimizer and the maximizer, respectively. The cost to optimize is $(x_2 - T)^2 + u_0^2 + u_1^2 - 2w_0^2 - 2w_1^2$. Our aim is to compute the feedback Nash equilibrium strategies over two time steps $t = 0, 1$, namely, $\sigma_t : \mathbb{R} \to \mathbb{R}$, for the minimizer, and $\gamma_t : \mathbb{R} \to \mathbb{R}$ for the maximizer.

a) We will use the dynamic programming approach. Given $V_2(x) = (x - T)^2$, complete the missing entries in the backward iteration for determining $V_1(x)$, namely, the cost-to-go at time 1. (1 point)

$$V_1(x) = \min_{u \in \mathbb{R}} \max_{?} \underbrace{\left[u^2 - 2w^2 + V_?(f(x, u, w))\right]}_{J(x,u,w)} \tag{1}$$

b) Now, using the dynamics, write the expression for $J(x, u, w)$. (1 point)

c) Observe that $J(x, u, w)$ is convex and differentiable in $u \in \mathbb{R}$ and concave and differentiable in $w \in \mathbb{R}$. Thus, explain how you would determine the $\min_{u\in\mathbb{R}} J(x, u, w)$ and the $\max_{w\in\mathbb{R}} J(x, u, w)$. (.5 point)

d) Compute the feedback Nash equilibrium strategies at time $t = 1$ putting the steps above together. You should obtain a pair of affine strategies $\sigma_1(x) = k_u x + b_u$ and $\gamma_1(x) = k_w x + b_w$. (2 points) *You may use the fact that the inverse of a* 2×2 *matrix is given as:*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}.$$

e) Using the strategies derived, compute the cost-to-go function $V_1(x)$. If you could not derive the explicit form of the strategies, you may substitute the affine strategies symbolically and continue. (1 point)

f) Now, write the backward iteration for computing $V_0(x)$. Would $\sigma_0(x), \gamma_0(x)$ be also affine? If so, would they have the same linear and offset terms (k_u, k_w, b_u, b_w)? (1.5 point)

Solution:

a) The missing entries are $w \in \mathbb{R}$ and 2:

$$V_1(x) = \min_{u\in\mathbb{R}} \max_{w\in\mathbb{R}} \left[u^2 - 2w^2 + V_2(f(x, u, w)) \right].$$

b) The expression of $J(x, u, w)$ is

$$J(x, u, w) = u^2 - 2w^2 + \left(\frac{1}{2}x + u + w - T \right)^2.$$

c) To determine the minimum and the maximum, we can take the derivatives of the cost function with respect to the two agents control actions and set it equal to zero.

d) To compute the feedback Nash equilibrium, first we take the derivatives of the cost function with respect to the two agents control actions and we set it equal to zero:

$$\begin{aligned} \frac{\partial J(x_0, u_0, u_1, w_0, w_1)}{\partial u_1} &= 2(\frac{1}{2}x_1 + u_1 + w_1 - T) + 2u_1 = 0 \\ \frac{\partial J(x_0, u_0, u_1, w_0, w_1)}{\partial w_1} &= 2(\frac{1}{2}x_1 + u_1 + w_1 - T) - 4w_1 = 0. \end{aligned}$$

We can write them in a matrix form:

$$\begin{bmatrix} 4 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ T \end{bmatrix}$$

For the matrix to be invertible $ad - bc$ needs to be different from zero. In our case we have

$$ad - bc = 4 * (-2) - 2 * 2 = -12.$$

We can then compute u_1 and w_1:

$$\begin{bmatrix} u_1 \\ w_1 \end{bmatrix} = -\frac{1}{12} \begin{bmatrix} -2 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ T \end{bmatrix} = \frac{1}{12} \begin{bmatrix} 4 & -8 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ T \end{bmatrix}$$

Thus, the Nash equilibrium policies are:

$$\begin{aligned} u_1 &= \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ T \end{bmatrix}, \\ w_1 &= \begin{bmatrix} \frac{1}{6} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ T \end{bmatrix} \end{aligned}$$

e) The cost to go is

$$V_1(x) = \left(\frac{1}{2}x - \frac{1}{3}x + \frac{2}{3}T + \frac{1}{6}x - \frac{1}{3}T - T\right)^2 + \left(-\frac{1}{3}x + \frac{2}{3}T\right)^2 - \left(\frac{1}{6}x - \frac{1}{3}T\right)^2$$
$$= \frac{7}{36}(x - 2T)^2.$$

f) To compute $V_0(x)$ we need to solve

$$V_0(x) = \min_{u\in\mathbb{R}} \max_{w\in\mathbb{R}} \left[u^2 - 2w^2 + V_1\left(f(x,u,w)\right)\right]$$
$$= \min_{u\in\mathbb{R}} \max_{w\in\mathbb{R}} \left[u^2 - 2w^2 + \frac{7}{36}\left(\frac{1}{2}x + u + w - 2T\right)^2\right].$$

To compute the Nash equilibrium, we need to take the derivatives with respect to u and to w. We would obtain two equations linear in all the terms. Thus, the solution for u and w would still be linear. However, the values of k_u, k_w, b_u, b_w) are going to differ from the previous ones we computed.

Problem 3. Shortest path game (9 points + 1 bonus)

Alice and Bob play the following dynamic game: A token is moved along a directed graph $\mathcal{G}$ with nodes $\{s_1, s_2, \dots, s_6\}$ that represent the state of the game. At time $t = 0$, the token is placed at s_1. Edges are associated with costs, and we let $c_{ij} \in \mathbb{R}_{>0}$ denote the cost of the directed edge from s_i to s_j.

Alice and Bob have the same state-dependent action sets $\mathcal{U}_s = \mathcal{V}_s$. Namely, for each node s, we have $\mathcal{U}_s = \mathcal{V}_s = \mathcal{N}(s)$ where $\mathcal{N}(s)$ is the set of nodes v for which there is an edge from s to v. For odd i, the token transitions to the node determined by Alice's action $u \in \mathcal{U}_{s_i}$; then Alice incurs cost c_{ij}, and Bob incurs no cost. Conversely, for even i, the token transitions to the node determined by Bob's action $v \in \mathcal{V}_{s_i}$; then Bob incurs cost c_{ij}, and Alice incurs no cost. If the token is at s_6, it stays there and the cost is 0 for both players. The game ends at $t = 5$.

Figure 1 shows the graph $\mathcal{G}$. Both Alice and Bob are assumed to be cost minimizers.

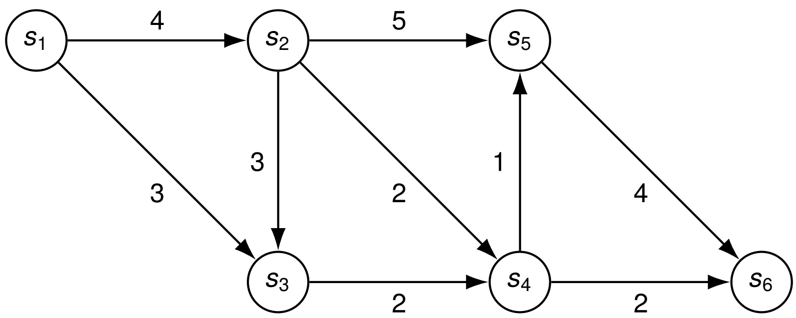


Figure 1

For $j \in \{1, 2, \dots, 6\}$ and $t \in \{0, 1, \dots, 5\}$, we define $V_t^A(s_j)$ as Alice's cost for the token to reach s_6 within at most t steps, assuming each player acts optimally with respect to her/his total cost). The respective cost incurred by Bob is $V_t^B(s_j)$. If s_6 is not reachable from s_j within at most t steps, we set the respective value to ∞.

a) Initializing $V_0^A(s) = 0$ for $s = s_6$ and $V_0^A(s) = \infty$ for $s \in \{s_1, \dots, s_5\}$, we can use dynamic programming to determine $V_t^A(s)$ for $t = 1, 2, \dots, 5$, and $s \in \{s_1, \dots, s_5\}$ as follows.

$$V_t^A(s_j) = \begin{cases} \min\limits_{s_k \in \mathcal{N}(s_j)} \left\{ c_{jk} + V_{t-1}^A(s_k) \right\}, & \text{if } j \text{ is odd;} \\ V_{t-1}^A(s_{k^\star}), & \text{otherwise, where } k^\star = \operatorname*{arg\,min}\limits_{s_k \in \mathcal{N}(s_j)} \left\{ c_{jk} + V_{t-1}^B(s_k) \right\}, \end{cases}$$

Write down the respective expression for determining $V_t^B(s_j)$. (1 point)

Solution:
We can write

$$V_t^A(s_j) = \begin{cases} \min\limits_{s_k \in \mathcal{N}(s_j)} \left\{ c_{jk} + V_{t-1}^A(s_k) \right\}, & \text{if } j \text{ is odd;} \\ V_{t-1}^A(s_{k^\star}), & \text{otherwise, where } k^\star = \operatorname*{arg\,min}\limits_{s_k \in \mathcal{N}(s_j)} \left\{ c_{jk} + V_{t-1}^B(s_k) \right\}, \end{cases}$$

and similarly,

$$V_t^B(s_j) = \begin{cases} \min\limits_{s_k \in \mathcal{N}(s_j)} \left\{ c_{jk} + V_{t-1}^B(s_k) \right\}, & \text{if } j \text{ is even;} \\ V_{t-1}^B(s_{k^\star}), & \text{otherwise, where } k^\star = \operatorname*{arg\,min}\limits_{s_k \in \mathcal{N}(s_j)} \left\{ c_{jk} + V_{t-1}^A(s_k) \right\}, \end{cases}$$

b) Let $u_t^\star(s)$ and $v_t^\star(s)$ be the action Alice and Bob must take in order to achieve cost $V_t^A(s)$ and $V_t^B(s)$, respectively. In Table 1 (next page), the box corresponding to state s and time t should contain $(V_t^A(s), V_t^B(s))$ at

s \ t	0	1	2	3	4	5
s_1	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	$(4,4)$ $(s_2,-)$	$(\quad,\quad)$ $(\quad,\quad)$	$(\quad,\quad)$ $(\quad,\quad)$
s_2	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	$(\quad,\quad)$ $(\quad,\quad)$	$(\quad,\quad)$ $(\quad,\quad)$	$(4,3)$ $(-,s_4)$	$(4,3)$ $(-,s_4)$
s_3	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	$(2,2)$ $(s_4,-)$	$(6,1)$ $(s_4,-)$	$(6,1)$ $(s_4,-)$	$(6,1)$ $(s_4,-)$
s_4	(∞,∞) $(-,-)$	$(\quad,\quad)$ $(\quad,\quad)$	$(\quad,\quad)$ $(\quad,\quad)$	$(4,1)$ $(-,s_5)$	$(4,1)$ $(-,s_5)$	$(4,1)$ $(-,s_5)$
s_5	(∞,∞) $(-,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$
s_6	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$

Table 1

the top, and $(u_t^*(s), v_t^*(s))$ at the bottom (where "$-$" means this player can choose any action). Fill out the 6 missing boxes. (6 points)

Solution:

s \ t	0	1	2	3	4	5
s_1	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	$(4,4)$ $(s_2,-)$	$(8,3)$ $(s_2,-)$	$(8,3)$ $(s_2,-)$
s_2	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	$(0,4)$ $(-,s_4)$	$(4,3)$ $(-,s_4)$	$(4,3)$ $(-,s_4)$	$(4,3)$ $(-,s_4)$
s_3	(∞,∞) $(-,-)$	(∞,∞) $(-,-)$	$(2,2)$ $(s_4,-)$	$(6,1)$ $(s_4,-)$	$(6,1)$ $(s_4,-)$	$(6,1)$ $(s_4,-)$
s_4	(∞,∞) $(-,-)$	$(0,2)$ $(-,s_6)$	$(4,1)$ $(-,s_5)$	$(4,1)$ $(-,s_5)$	$(4,1)$ $(-,s_5)$	$(4,1)$ $(-,s_5)$
s_5	(∞,∞) $(-,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$	$(4,0)$ $(s_6,-)$
s_6	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$	$(0,0)$ $(-,-)$

c) Write down a strategy γ_A for Alice and γ_B for Bob such that (γ_A, γ_B) is a subgame perfect equilibrium of the above game. *Hint: Look at the last column of the table above.* (2 points)

Solution:
Alice's strategy γ_A must satisfy

$$\gamma_A(s_1) = s_2,$$
$$\gamma_A(s_3) = s_4,$$
$$\gamma_A(s_5) = s_6,$$

and Bob's strategy γ_B must satisfy

$$\gamma_B(s_2) = s_4,$$
$$\gamma_B(s_4) = s_5.$$

The remaining values $\gamma_A(s_2), \gamma_A(s_4)$ and $\gamma_B(s_1), \gamma_B(s_3), \gamma_B(s_5)$ can be chosen arbitrarily, as these actions affect neither state transitions nor costs.

d) (bonus) Now instead of having players minimize their own cost, suppose they aim to minimize the sum of both their total costs, i.e. the social cost. Write down a strategy $\widehat{\gamma}_A$ for Alice and $\widehat{\gamma}_B$ for Bob such that $(\widehat{\gamma}_A, \widehat{\gamma}_B)$ is a subgame perfect equilibrium of this modified game. *Hint: The strategies can be inferred directly by looking at the graph.* How does the social cost of $(\widehat{\gamma}_A, \widehat{\gamma}_B)$ compare to that of (γ_A, γ_B)? (1 point)

Solution:
In this case, any pair of strategies that leads to choosing the overall shortest path from each node to s_6 constitutes a subgame perfect equilibrium. Such strategies must satisfy

$$\widehat{\gamma}_A(s_1) = s_3,$$
$$\widehat{\gamma}_A(s_3) = s_4,$$
$$\widehat{\gamma}_A(s_5) = s_6,$$

and

$$\widehat{\gamma}_B(s_2) = s_4,$$
$$\widehat{\gamma}_B(s_4) = s_6.$$

The remaining values $\widehat{\gamma}_A(s_2), \widehat{\gamma}_A(s_4)$ and $\widehat{\gamma}_B(s_1), \widehat{\gamma}_B(s_3), \widehat{\gamma}_B(s_5)$ can be chosen arbitrarily, as these actions affect neither state transitions nor costs. The social cost of $(\widehat{\gamma}_A, \widehat{\gamma}_B)$ is 7 and thus strictly lower than the social cost of (γ_A, γ_B) which is 11.