IEOR@IITB

*IE 708 Markov Decision Processes*
Class Test 1 (30/Aug/2024)

*Attempt all questions. If you feel the need to use results not discussed in the class, please state them clearly. State all assumptions and arguments clearly.*

---

1. A firm wants to announce a one time discount offer for a single product in a price sensitive market. Model this as a MDP, i.e., list the elements of the 5 tuple that define a MDP (state space, action space, reward functions, transition probabilities and the planning horizon).

   **Soln.** (state space (1), action space (0.5), reward functions (1), transition probabilities (1.5) and the planning horizon (0.5))

   State space ($\mathcal{S}$): The states can represent the current stock of the product as well as if it is currently the discount phase or pre-discount phase. So a state $s \in \mathcal{S}$ can be a two tuple, $s = (k, l)$, where $k \in \{0, 1, 2, ..\}$ is the current product stock and $l \in \{0, 1\}$ denotes the pre-discount(0)/discounted(1) phase.

   Action space ($\mathcal{A}$): The pre-discount phase states can have at least two actions, i.e., to announce or not the discount. More than two actions can be considered for deciding the discount amount (10%, 20%, etc). Discounted phase states will not have any effective actions.

   Reward function ($r_t(s, a, s')$): Should reflect sales/profit achieved by the firm in the given state based on the random demand which depends on the price set by the firm via available actions.

   Transition probabilities ($p_t(s'|s, a)$): The discounted phase states can't transition to any pre-discounted phase states. If you assume no additional inventory is added, states can only transition to other states with lesser stock. The transition probabilities should appropriately depend on the current price, the difference in stocks $s - s'$ and random demand.

   Planning horizon: Can be finite for $M$ months $t \in \{0, 1, 2, \ldots, M\}$.

   See section 2.1 in `https://github.com/martyput/MDP_book/blob/6ffd0b961ad2dcc7813421a59800a` `Chapter-3/Book/Chapter%203%20July%202023.pdf` for reference.

2. Consider the a shortest past problem illustrated in Figure 1.

   (a) Argue why the the straight forward way of exploring the entire feasible solution space of all paths and their costs is usually not pursured. **Soln.** (0.5) For a dense graph with $n$ vertices, the feasible space of all paths will be in the order of $n!$.

   (b) Use a dynamic programming, DP, recursion to find the optimal path. **Soln.** (1.5) See the Floyd-Warshall algorithm or another DP based algo.

   (c) Using the above, find the optimal path between B and H. **Soln.** (0.5) $B, F, H$. Cost $= 10$

   (d) What is the minimal cost between C and H? Note that while the optimal cost/value is unique, the optimal path is not. **Soln.** (1) Cost $= 9$. Path - and $C, D, G, H$.

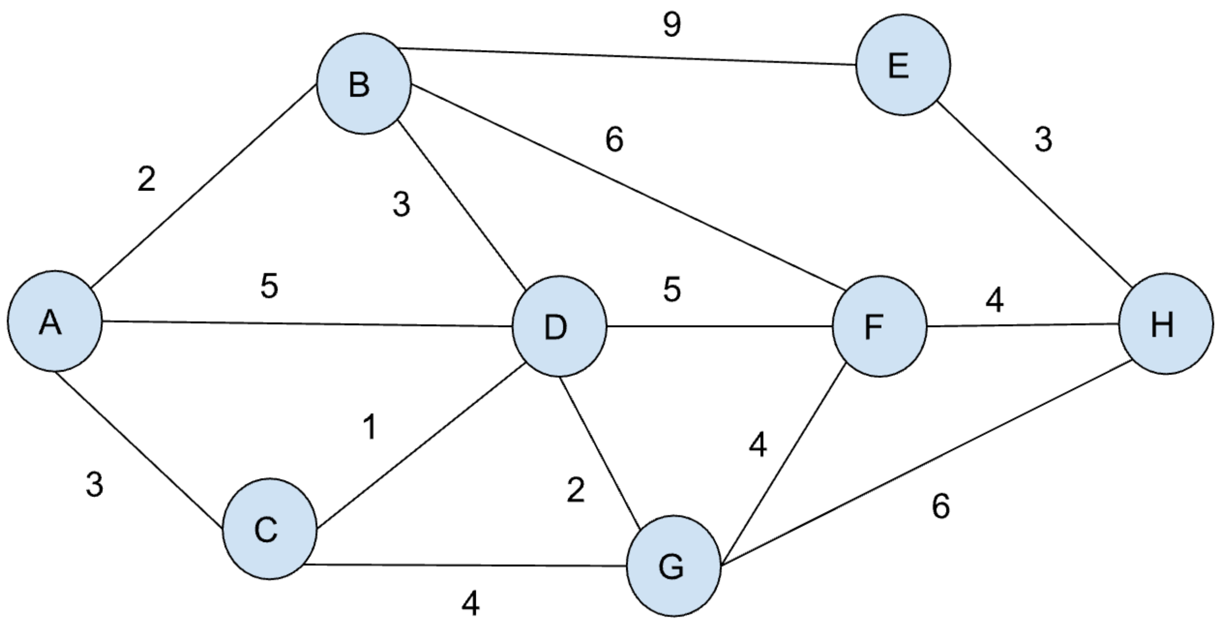


Figure 1: An instance of SPP (shortest path problem)

3. Check/verify which of the following (real valued) functions are superadditive, subadditive or neither?

   (a) $g_1(x, y) = h(x + y)$ for a convex increasing function $h(\cdot)$ on $\mathbf{R}$.
   **Soln.** (1) let $t_1 \in (0, 1)$ be such that $t_1(x^+ + y^+) + (1 - t_1)(x^- + y^-) = x^+ + y^-$. Therefore, $t_1 = \frac{x^+ - x^-}{x^+ + y^+ - x^- - y^-}$. Similarly, for $t_2 \in (0, 1)$ such that $t_2(x^+ + y^+) + (1 - t_2)(x^- + y^-) = x^- + y^+$, we get $t_2 = \frac{y^+ - y^-}{x^+ + y^+ - x^- - y^-}$. Therefore, $t_1 = 1 - t_2$
   As $h$ is convex,

$$t_1 h(x^+ + y^+) + (1 - t_1)h(x^- + y^-) \geq h(x^+ + y^-) \tag{1}$$

$$t_2 h(x^+ + y^+) + (1 - t_2)h(x^- + y^-) \geq h(x^- + y^+) \tag{2}$$

   Adding (1) and (2) superadditivity is satisfied.

   (b) $g_2(x, y) = xy$ for reals $x$ and $y$. **Soln.** (1)

$$x^+ \geq x^-, y^+ \geq y^-$$

$$(x^+ - x^-)(y^+ - y^-) \geq 0$$

$$x^+y^+ - x^+y^- \geq x^-y^+ - x^-y^-$$

   $g_2$ is superadditive.