VPP-A exam, 09.12.2025

Problem 1 [20b]

Program the DP algorithm for the following problem:

- System equation: $x_{k+1} = \min\{\max\{x_k + 2 \cdot u_k - 2 \cdot w_k - 3, 1\}, c\}, \quad k = 0, 1, \ldots, 19 \quad (i.e., N = 20)$
- Feasible states and control: $x_k \in \{1, \ldots, c\}, \quad u_k \in \{1, \ldots, d\}$ (natural numbers only)
- Random variable w can attain six values $w \in \{1, \ldots, 6\}$ with equal probability.
- kth stage cost: $g(x_k, u_k, w_k) = x_k + 2 \cdot u_k - 5 \cdot \min\{x_k, 5\}$
- Final stage cost: $g_N(x_N) = 3 \cdot x_N + \min\{x_N, 3\}$
- $J_k^*(x_k) = \min_{\pi^k} E_{w_k,\ldots,w_{N-1}} \left\{ g_N(x_N) + \sum_{i=k}^{N-1} g_i(x_i, u_i, w_i) \right\}$
- Display $J(x_0)$ for $x_0 = 2, c = 12, d = 8$.

Problem 2 [20b]

Program the value iteration algorithm [10b] and policy iteration algorithm [10b] for the following problem:

- Markov chain, where the transition to the next state is governed by the equation

  $$x_{k+1} = w_k,$$

  where w_k is a random variable with probability distribution: $P\{w_k = j \,|\, x_k = i, u_k = u\} = p_{ij}(u)$ (dependend on x_k, u_k), the matrix of transition probabilities can be found in `ex_2025_12_09_P.mat` (in the format $P(x_k, x_{k+1}, u_k)$)
- Cost function: $g(x, u) = 40 \cdot x + u^2$
- Discount factor: $\alpha = 0.99$
- Feasible states and control: $x \in [1, 2, \ldots, 9], \quad u \in [1, 2, \ldots, 10]$
- Display $J^*(x)$ and $\mu^*(x)$ for all states, with

  $$J^*(i) = \min_{\mu} \lim_{N \to \infty} E \left\{ \sum_{k=0}^{N-1} \alpha^k g(x_k, \mu(x_k)) \,|\, x_0 = i \right\}$$

- Criterion for convergence (termination of the while loop): $\max(|J_{k+1} - J_k|) < 10^{-5}$