Dynamic Programming & Stochastic Control (SC42110)
Exercise Set 2

Thursday 22^{nd} May, 2025

Exercise 1 (Matlab exercise)**.** Consider the machine replacement problem of Example 3.3.

a) Write the Matlab code of DPA for solving this problem. The code must output the optimal costs-to-go $(J_t)_{t=0}^{T}$ and the optimal policy $(\mu_t)_{t=0}^{T-1}$.

b) Run $N = 10$ simulations of the system under the optimal policy for the initial state $X_0 = 1$ and compute the average (over runs) cost of the trajectories.

c) Run $N = 10$ simulations of the system under the sub-optimal policy $(\nu_t)_{t=0}^{T-1}$ where for all $t \in \{0, \dots, T-1\}$,

$$\nu_t(x) = \begin{cases} \text{wait} & \text{if } x \in \{2, 3\}, \\ \mu_t(x) & \text{otherwise.} \end{cases}$$

Compute the average (over runs) cost of the trajectories and discuss the result in comparison with that of part (c).

d) Repeat parts (b) and (c) with $N = 1000$.

Exercise 2. Consider a random process X_k that obeys the difference equation

$$\begin{cases} X_{k+1} = aU_k + (1-a)W_{k+1}, & k = 0, 1, \\ X_0 = -1. \end{cases}$$

where $a \in [0, 1]$ is an scalar, $U_k \in \mathbb{U} = [0, 1]$ is the input, and $W_{k+1} \in \mathbb{W} = \{0, 1\}$ is the disturbance such that $W_{k+1} = 0$ if $X_k < 0$ and W_{k+1} takes values 0 and 1 with equal probability if $X_k \geq 0$. Find the minimum of the cost function

$$\mathbb{E}\left(X_2^2 + \sum_{k=0}^{1} (X_k^2 + U_k^2) \right)$$

and an optimal control policy $U_k = \mu_k(X_k) \in \mathbb{U},\ k = 0, 1$.

Exercise 3. Consider a machine that is either running or is broken down. Suppose that:

- If the machine runs throughout the week, it makes a profit of $G > 0$ for that week;
- If the machine fails during the week, the profit is zero for that week.

At the beginning of each week, we have to decide what to do with the machine. When the machine is running at the start of the week, we have the following options:

- We can perform preventive maintenance. The probability that the machine will run throughout the week after this maintenance is p_m. Notice that the maintenance will cost $C_m > 0$;
- We can choose to not perform any maintenance. The probability of running throughout the week is then $p_{nm} < p_m$.

On the other hand, when the machine is broken at the start of the week, we have the following options:

- The machine can be repaired at a cost of $C_r > C_m$. It will then run throughout the week with probability p_m;
- The machine can be replaced at a cost of $C_l > C_r$ by a new machine. Now, it is guaranteed that the machine will run throughout the week.

Notice that it is not allowed to simply leave the machine broken. After $N > 1$ weeks, the machine, irrespective of its state, is scrapped without incurring any cost. Our goal is to find the optimal policy for each week (to perform the maintenance or not, to repair or to replace). The questions now follow.

a) Show how this problem can be written in the standard form, i.e., introduce the state space, define the set of admissible controls, write the equations of the system's dynamics and <u>reward</u> function.

b) Find the optimal policy for the N-th and $(N-1)$-th weeks.

Exercise 4. There are n tickets available to an event. We plan to sell these tickets during N days. Suppose that on each day i,

- the price $p_i > 0$ for tickets is fixed;
- we decide on the number of tickets that we plan to sell at the beginning of this day;
- the demand D_i for the number of tickets is a random number between 0 and n with a known probability distribution;
- there is no penalty for unsold tickets.

Our goal is now to maximize the expected revenue and find the corresponding optimal strategy. The questions now follow.

a) Formulate the corresponding dynamic programming problem in the standard form, i.e., define the state space, the control space, the system's dynamics, and the cost function to be optimized. (*Hint:* The control action has state-dependent constraints.)

b) Suppose now $n = 6$, $N = 2$, $p_1 = 20$, $p_2 = 40$, $D_1 = 6$, and

$$D_2 = \begin{cases} 3 & \text{with probability } 0.25, \\ 4 & \text{with probability } 0.5, \\ 5 & \text{with probability } 0.25. \end{cases}$$

(It is not difficult to see that we are guaranteed to sell all tickets we put for sale on the first day.) Find the maximum expected revenue and the corresponding optimal selling strategy in this special situation.

Exercise 5. Consider the dynamical system

$$X_{k+1} = X_k + U_k + W_{k+1}, \quad k = 0, 1, 2, 3,$$

where the input signal is $U_k \in \mathbb{U} = \{-1, 0, 1\}$ and the disturbance sequence $(W_k)_{k=1}^4$ is generated by a sequence of i.i.d. random variables such that $W_k \in \{-1, 1\}$ with an equal probability. Furthermore, we require the *state constraint* $-2 \leq X_k \leq 4$ to hold for all $k \in \{1, 2, 3, 4\}$, and any control policy should guarantee this, despite the random disturbance.

Find the minimal value of the cost functional

$$\mathbb{E}\left(\sum_{k=0}^{3} g(X_k, U_k)\right),$$

staring from initial state $X_0 = 0$, where

$$g(x, u) = x^2 + u^2 + a(x), \qquad a(x) = \begin{cases} 0 & \text{if } x \in [0, 2], \\ 13 & \text{otherwise.} \end{cases}$$

Exercise 6. Consider the dynamical system

$$X_{k+1} = X_k + U_k + W_{k+1}, \quad k = 0, 1, \ldots,$$

where the control input U_k is unconstrained, the disturbance $W_k \in [-1, 1]$ is i.i.d. with uniform distribution (i.e., the probability density of W_k is $1/2$ on $[-1, 1]$ and 0 beyond it). The cost functional is a mixture of linear and quadratic terms given by

$$J_0(x) = \mathbb{E}\left(X_2 + U_1^2 + U_0^2 \mid X_0 = x\right).$$

Find the optimal (minimal) cost and its corresponding optimal control policy $U_k = \mu_k(X_k)$, $k = 0, 1$.

Exercise 7. You own an oil well from which oil can be extracted at a rate of 100 barrels per year at a cost of \$85 per barrel. The oil price for the current year is \$100 per barrel, but the future oil prices are subject to uncertainty. Each year the price either increases or decreases by 10% with equal probabilities. (We assume that the oil price remains unchanged throughout the year). At the beginning of each year, you can decide whether to extract oil during that year.

a) Calculate the maximum expected profit that you can make during a period of three years.

b) Suppose now that you have an additional option of installing a new machine offered by a partner company. The machine can enhance your operation by increasing the maximum extraction capability from 100 barrels per year to 140 barrels per year, but it would increase an extraction cost from \$85 to \$91 per barrel. This new machine can be deployed at the beginning of any year, and once in place, it applies to all future years. Calculate again the maximum expected profit that you can make during a period of three years and determine when the new machine should be installed.