Dynamic Programming & Stochastic Control (SC42110)
Exercise Set 3

Monday 9$^{\text{th}}$ June, 2025

Exercise 1 (Linear quadratic control)**.** Prove Lemma 4.1. (*Hint:* Notice the assumptions on the disturbance and be mindful of matrix calculations. You may use the fact that for a random vector $W \in \mathbf{R}^n$ and a square matrix $K \in \mathbf{R}^{n\times n}$, we have $\mathbb{E}(W^\top K W) = \operatorname{tr}\big(K \operatorname{cov}(W)\big)$, where $\operatorname{cov}(W)$ is the covariance matrix of W and $\operatorname{tr}(M) = \sum_{i\in[n]} M(i,i)$ is the trace of the matrix $M \in \mathbf{R}^{n\times n}$.)

Exercise 2 (LQ control with random dynamics*)**.** Consider the linear dynamics

$$X_{t+1} = A_{t+1}X_t + B_{t+1}U_t, \quad t = 0, 1, \ldots,$$

where $X_t, U_t \in \mathbf{R}$ are the state and action variables, respectively. Assume that

- $A_t \in \mathbf{R}$ is an i.i.d. random variable with $\mathbb{E}(A_t) = a_1 \neq 0$ and $\mathbb{E}(A_t^2) = a_2$;
- $B_t \in \mathbf{R}$ is an i.i.d. random variable with $\mathbb{E}(B_t) = b_1 \neq 0$ and $\mathbb{E}(B_t^2) = b_2$.

Consider the optimal control problem

$$\min_{U_0,\ldots,U_{T-1}} \mathbb{E}\left\{\sum_{t=0}^{T-1}(X_t^2 + U_t^2) + X_T^2 \;\middle|\; X_0 = x\right\},$$

over the planning horizon $T \in \mathbf{N}$ with initial state $x \in \mathbf{R}$.

a) Show that the optimal cost-to-go is of the form $J_t(x) = k_t x^2$ with $k_t > 0$ and the optimal policy is of the form $\mu_t(x) = \ell_t x$ for each t. Derive a recursive formula for the parameters k_t and ℓ_t. (*Hint:* Treat A_t and B_t as disturbance and note that $a_2 - a_1^2 = \operatorname{var}(A_t) \geq 0$ and similarly $b_2 \geq b_1^2$.)

b) Does the optimal policy converge to a stationary one as $T \to \infty$? Why?

c) Does certainty equivalence hold for this problem? Why?

Exercise 3 (Optimal portfolio with logarithmic utility)**.** Prove Lemma 4.6.

Exercise 4 (Optimal purchase with deadline)**.** Assume that a fixed quantity of some raw material is needed by a deadline. We can decide to buy at today's price or to wait a period during which the price can go up or down due to the fluctuations of the market. The purchase must be made within T *time periods*. Let us assume that the successive prices $W_t \in \mathbf{R}_+$, $t \in \{1, \ldots, T\}$, of the material are i.i.d with a given p.d.f. $p_{W_t}(w) = \rho(w)$ for $w \in \mathbf{R}_+$. Our goal is to *minimize the expected price of the purchase.*

a) Formulate this problem as an optimal stopping problem by defining the state and action spaces, the dynamics (and the disturbance if needed), and the costs.

b) Show that the optimal purchase policy is in the form of a threshold, that is, the optimal decision is to buy at the price W_t if $W_t \leq \alpha_t$ and to wait otherwise.

c) Find α_{T-1} and a recursive formula $\alpha_t = \mathcal{F}(\alpha_{t+1})$ for the purchase thresholds.

d) Show that the purchase threshold increases as the deadline approaches.

Exercise 5. Peyman is leaving his university. His friends are planning to organize a series of farewell parties for him every day, from Monday (1st day) till Friday (5th day). There is a pot of money that is used for both organizing the parties and Peyman's farewell gift. The university has contributed €100 to the pot. Let X_{t-1} be the amount of money in the pot *at the beginning of* t-th day with $X_0 = 100$. Peyman has two options *at the beginning of* each t-th day:

- He can *bail*, in which case, he will receive all the money in the pot as his farewell gift and there will be no party on that day and afterward.
- If he does *not bail*, 60% of the pot is spent to organize the party on the same day and a random amount of money $W_t > 0$ will be added to the pot by Peyman's friends after the party, where W_t is i.i.d. with a known distribution.

If there is a party on day 5, Peyman will receive the remaining money in the pot as his farewell gift after the party. Peyman aims to use dynamic programming to maximize the expected value of his farewell gift.

a) Formulate the corresponding DP problem by defining the state and control spaces, the system's dynamics with its disturbance, and the reward functions.

c) Use dynamic programming to find the optimal policy for Peyman on the day t assuming that the value function J_{t+1} is known.

b) Compute the value function and the corresponding optimal policy for bailing at the beginning of the last day.

c) Show that that J_t is *positive* and *convex* for all $x > 0$ and all $t = 0, \ldots, T$.[1]

[1]A function $J : \mathbb{X} \to \mathbf{R}$ is *convex* if and only if for any $\alpha \in [0, 1]$ and any $x, y \in \mathbb{X}$, $J(\alpha x + (1 - \alpha)y) \leq \alpha J(x) + (1 - \alpha)J(y)$.

Exercise 6. A *call* (respectively, *put*) *option* is a financial contract that gives the owner the right (but *not* the obligation) to buy (respectively, to sell) a prescribed stock at a specified price, called the *strike price* K, at any time on or before an *expiration date* T. An option can be exercised at most once during its lifetime, and we suppose that the option can be exercised only at the end of a month. Denote by X_t the price of the underlying stock at the end of the t-th month. If a call option is exercised at the end of the t-th month, the owner earns an immediate profit of $X_t - K$ by buying the prescribed stock at the strike price K and selling it immediately on the market at price X_t. Similarly, exercising a put option earns its owner an immediate profit of $K - X_t$. When the option is exercised, the immediate profit will be invested at a fixed (monthly) interest rate r until time T. Assume that the current price of the stock is $X_0 = \$100$. During any one-month period, the stock can either increase by a factor of 1.2 with probability $\frac{2}{3}$ or decrease by a factor of 0.9 with probability $\frac{1}{3}$. Assume further that r is equal to 10%. Determine the optimal exercise policies for the options below, assuming that the owner aims to maximize expected profits.

a) A call option with a strike price \$105 and an expiration date in 2 months from now.

b) A put option with a strike price \$105 and an expiration date in 2 months from now.

Exercise 7 (Batcoin*)**.** A speculator owns *one Batcoin* (Gotham City's new cryptocurrency) which can be thought of as a financial instrument. Each day, she has an opportunity to sell this Batcoin, which she may either accept or reject. The potential sale prices Z_t for day t are independently and identically distributed with probability density function $\rho(z) = 2z^{-3},\ z \geq 1$. As Batcoins represent an extremely risky investment, each day there is a probability $1 - \beta$ that the market for Batcoins will collapse, making her Batcoin completely worthless. Finally, we assume that the speculator has to sell her Batcoin within a period of T days, that her Batcoin can only be sold as a whole, and that her objective is to maximize the *expected earning* from selling this Batcoin.

a) Formulate a dynamic program to solve this problem. Then, argue that the speculator should sell her Batcoin at day $t \in \{1, \dots, T\}$ if $Z_t \geq \alpha_t$ for some constant α_t (provided that she has not done so yet and that the market has not yet collapsed).

b) Part (a) implies that the optimal policy for the speculator is completely determined by the constants $\alpha_t,\ t = 1, \dots, T$. Show that these constants can be determined from a recursion, that is, find a function F such that $\alpha_t = F(\alpha_{t+1})$.

c) Assume that $\beta \leq \frac{1}{2}$ (that is, the Batcoin market has a smaller chance of surviving than collapsing). Argue that it is optimal for the speculator to sell her Batcoin right away.

d) Assume otherwise that $\beta > \frac{1}{2}$. If the speculator is a long-term investor (that is, $T \to \infty$), then the sequence $\alpha_T, \alpha_{T-1}, \dots, \alpha_1$ converges to a fixed point of the function F, namely $\alpha^\star = F(\alpha^\star)$. In this case, her (approximately) optimal policy is to keep hold of her Batcoin unless its price exceeds $\alpha^\star$. Determine $\alpha^\star$.