LINMA2222 | Project

Part 1: Modeling and Analysis

Instructions

This project is due on **October 10, 2025, at 11:59 pm**. You may form teams of up to two people. Each team should upload a single zip-file, containing their report and code, on Moodle. Make sure to clearly state the team members in your report. You can do the project in MATLAB *or* Python *or* Julia, but not a mix.

1 Introduction

This project studies a *portfolio optimization problem*. Given a financial asset a, with price p_t at time t, the goal is to determine an optimal strategy for buying or selling units of a, based on the current portfolio state and available information, at each discrete time step $t = 0, 1, 2, \ldots$.

2 The model

Let $q_t \in \mathbb{R}$ denote our *inventory* at time t, that is, the quantity of asset a (e.g., a stock, gold, etc.) we hold at time t, expressed as a fraction of the total market supply (e.g., 1000 stocks/units). For simplicity, we assume that the decision of buying or selling inventory is taken only at discrete time intervals (e.g., at the beginning or end of each day); hence, we let $t \in \mathbb{N}$.

Let $p_t \in \mathbb{R}$ be the *market price* at time t (this is the price that we observe and we use to make decisions), and let $\bar{p}_{t,t+1} \in \mathbb{R}$ be the average execution price on the interval $[t, t+1)$ (this is the price at which the trade actually takes place, which may differ from p_t).

We consider the following *gross* stage reward, which measures our profit made from trading:

$$g_t \coloneqq 1000 \cdot (q_{t+1}p_{t+1} - q_t p_t - (q_{t+1} - q_t)\bar{p}_{t,t+1}).$$

Question 2.1 Provide an interpretation of g_t. What does it represent financially?

Given a *utility function* $c : \mathbb{R} \to \mathbb{R}$, we consider the following stage reward (sometimes called *net* or *utile* stage reward):

$$r_t = c(g_t).$$

The utility function allows to capture different phenomena, like taxes, aversion to risk, etc.

The variable we can control, hereafter called control input, is denoted by $u_t \in \mathbb{R}$ and represents the quantity of the asset a (again as a fraction of the asset's total availability) that we buy ($u_t \geq 0$) or sell ($u_t \leq 0$) at time $t \in \mathbb{N}$. Hence, for all $t \in \mathbb{N}$,

$$q_{t+1} = q_t + u_t.$$

Objective The goal of a trader is to find policies that maximize the infinite-horizon average reward:

$$\lim_{T\to\infty} \frac{1}{T} \mathbb{E}\Big[\sum_{t=0}^{T-1} r_t\Big]. \tag{1}$$

The admissible policies take the form $u_t = \pi(x_t)$ (if deterministic) or $u_t \sim \pi(\cdot \mid x_t)$ (if randomized), where x_t is the vector of system state variables, described in the model below.

Model We next describe the model that governs the evolution of the system. The state vector is $x_t = [q_t, z_t^a, z_t^u]^\top$, where q_t is the inventory as above, and z_t^a and z_t^u are two other variables modeling components (one stochastic and one input-dependent) of the variation of the price p_t, and described next:

$\theta = 0.5$	$\omega^a = 0.1$	$\omega^u = 0.2$
$\sigma^a = 0.018$	$\beta^u = -0.048$	$\gamma^u = 0.06$
	$\sigma^p = 0.02$	

Table 1: Model parameters.

- The evolution of z_t^a is given by

$$z_{t+1}^a = (1-\omega^a)z_t^a + \omega^a\sigma^a\xi_t^a, \quad \xi_t^a \sim \mathcal{N}(0,1),$$

  where $\omega^a \in (0,1)$ and $\sigma^a \in \mathbb{R}_{\geq 0}$ are parameters.
- The evolution of z_t^u is given by

$$z_{t+1}^u = (1-\omega^u)z_t^u + \omega^u\beta^u u_t,$$

  where $\omega^u \in (0,1)$ and $\beta^u \in \mathbb{R}$ are parameters.
- The evolution of the price p_t is given by

$$p_{t+1} = p_t + z_t^a + z_t^u + \gamma^u u_t + \sigma^p\xi_t^p, \quad \xi_t^p \sim \mathcal{N}(0,1),$$

  where $\gamma^u \in \mathbb{R}$ and $\sigma^p \in \mathbb{R}_{\geq 0}$ are parameters.
- The execution price $\bar{p}_{t,t+1}$ is given by

$$\bar{p}_{t,t+1} = \theta p_t + (1-\theta)p_{t+1} + \theta\gamma^u u_t,$$

  where $\theta \in (0,1)$ is a parameter.

We stress that, in the above dynamics, $\{\xi_t^a\}_{t\in\mathbb{N}}$ and $\{\xi_t^p\}_{t\in\mathbb{N}}$ are independent random variables with standard normal distribution. We let $\xi_t = [\xi_t^a, \xi_t^p]^\top$ for all $t \in \mathbb{N}$.

Question 2.2 Provide an interpretation of the model and explain why it offers a reasonable representation of the evolution of an asset's price over time.

Question 2.3 Express the gross stage reward g_t as a function of x_t, u_t and ξ_t. In particular, compute explicitly the matrix H such that $g_t = \frac{1}{2}[x_t^\top, u_t^\top, \xi_t^\top]H[x_t^\top, u_t^\top, \xi_t^\top]^\top$.

Utility function Finally, we focus on the following instance for the utility function:

$$c(g) = \max\left(g - \frac{1}{2}g^2, 1 - \exp(-g)\right). \tag{2}$$

Question 2.4 Show that the utility function above is penalizing stage rewards g_t with a large variance. More precisely, suppose g_t is a random variable with some distribution of your choice (e.g., normal, uniform, etc.) with zero mean and variance σ^2; show empirically and/or analytically that $\mathbb{E}[c(g_t)]$ is a decreasing function of σ^2.

3 Closed-loop analysis

In this subsection, we use the parameters given in Table 1.

3.1 Unconstrained case

We consider the policy $\pi_{\text{cl}}(x_t) = K_{\text{cl}}x_t$, where $K_{\text{cl}} = [-0.5, 0.5, 0.5]$.

Question 3.1 Plot a sample trajectory of the closed-loop system with policy π_{cl}, starting at $x_0 = 0$, with horizon $t = 1, \ldots, T$, $T = 1000$. In your plots, show the variables x_t and u_t, as well as the average reward $\frac{1}{t}\sum_{s=0}^{t-1} r_s$, for $t = 0, 1, \ldots, T$. Interpret the results.

Question 3.2 Generate $N = 100$ sample trajectories, each of length $T = 1000$, of the closed-loop system with policy π_{cl}, starting at $x_0 = 0$. Average over the N sample trajectories and, for the averaged quantities, 1) show in a plot the mean and variance, 2) compute the average reward $\frac{1}{t}\sum_{s=0}^{t-1} r_s$, $t = 0, 1, \ldots, T$, and display it in a plot. Interpret the results.

3.2 Constrained case

The policy considered in the previous subsection may be infeasible, since it could lead to an inventory q_t that is smaller than zero (meaning that the investor holds a negative amount of the assets) or larger than one (meaning that the investor holds more than 100% of the market availability). Therefore, we next consider a *projected policy*, given by

$$\pi_{\mathrm{pcl}}(x_t) = \max(-q_t, \min(1 - q_t, \pi_{\mathrm{cl}}(x_t))).$$

Question 3.3 Repeat question 3.1 for π_{pcl}.

Question 3.4 Repeat question 3.2 for π_{pcl}.

3.3 A better policy?

Finally, propose a policy π_{yours} that satisfies the constraints on q_t ($0 \leq q_t \leq 1$) and potentially provides a better infinite-horizon average reward. You may use any reasoning you prefer (your investor intuition, a mathematical model, or otherwise), provided that it is clearly justified.

Question 3.5 Describe and motivate *briefly* the policy π_{yours}. Show that it does better than the proposed policy (if so!).

Question 3.6 Repeat question 3.1 for π_{yours}.

Question 3.7 Repeat question 3.2 for π_{yours}.