Summer 2025

DDQN to Optimize the Trading of a Double Mean-Reverting Asset

Based on the ideas provided from https://github.com/sebjai/ddpg-stat-arb-arxiv.git (1).

As seen in (1), a DDPG RL algorithm is used to optimize the trading of a mean-reverting asset (an asset that follows an OU-process). In particular, the original problem setup is the trading of an asset denoted by S_t that follows the SDE

$$dS_t = \kappa(\theta_A - S_t) + \sigma dB_t$$

discretizing time, the reward received by an agent for taking an action $a_t = I_{t+1}$ (inventory of the asset in the next period) is given by

$$r_{t+1} = a_t(S_{t+1} - S_t) - \lambda|a_t - I_t|S_t$$

where λ is the transaction costs incurred by the agent. What changes about optimal actions when parameters of this process change? What happens when a different RL algorithm is used?

Problem 1: Exploration

Before tackling these questions, we first recommend you play around with the code sent by Prof. Jaimungal. See how the DDPG algorithm is implemented and understand the environment class. Make sure you understand the loss function and the use of target networks.

Problem 2: DDQN vs DDPG

- Explain the main differences between DDQN and DDPG
- Outline some examples where DDPG would be favourable in comparison to DDQN? And vice versa?

Problem 3: DDQN

- Implement DDQN into the original problem provided in (1)
- Play around with the parameters λ, σ, κ, how do your heatmaps change?

Problem 4: Double Mean-Revert

Consider the following

$$dS_t = \kappa(\theta_A - S_t)\mathbb{I}_{S_t \leq \theta} + \kappa(\theta_B - S_t)\mathbb{I}_{S_t > \theta} + \sigma dB_t$$

where $\theta_A < \theta_B$ and $\theta = \frac{\theta_A + \theta_B}{2}$. The above SDE has the solution

$$S_{t+\Delta t} = \theta_A \left(1 - e^{-\kappa\,\Delta t}\right) \mathbb{I}_{S_t > \theta} + \theta_B \left(1 - e^{-\kappa\,\Delta t}\right) \mathbb{I}_{S_t \leq \theta} + S_t\, e^{-\kappa\,\Delta t} + \sigma \sqrt{\frac{1 - e^{2\kappa\,\Delta t}}{2\kappa}} B_{t+\Delta t}.$$

- Implement this asset price process and plot out simulations to verify its accuracy. Play with $\theta_{A,B}$, κ, and σ to see how these impact the dynamics.
- Update the environment in Problem 2 to use this process for the asset price and report what you find. Perform the same sensitivity analysis as was done in Problem 2.

References

[1] Sebastian Jaimungal. Ddpg-stat-arb-arxiv. https://github.com/sebjai/ddpg-stat-arb-arxiv.git. Accessed: 2025-05-30.