Hedging option with Actor Critics methods

NTHU DRL Final Project

110062301

資工大三 楊晨鍾

1. Abstract

In this project, I implement a RL model using DDPG to hedge the potential risk of selling a European call option by trading the underlying asset. I used adjusted closed apple stock price from 2012 to 2024 from yahoo finance as my data, the first 80% of the data is for training and the rest are for testing. It reached the same mean price and smaller standard deviation compared to delta hedging using Black-Scholes model.

2. Introduction

a. **European call option**

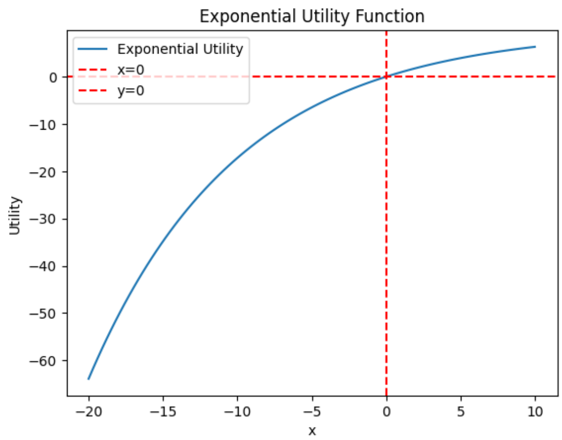
   i. Options are financial derivatives that provide the holder with the right, but not the obligation, to buy or sell an underlying asset at a specified price on or before a specified date.

   ii. A European call option is a type of option that can only be exercised at expiration. It allows the holder to buy the underlying asset at a predetermined strike price.

   iii. The seller of the call option faces potential risk if the underlying asset's price increases significantly.

b. **Risk-Averse and Value Function**:

i. Since the sadness of losing 10 dollars is not equal to the happiness of earning 10 dollars, we need a function to adjust the money. E.g. exponential utility



ii. The goal is to maximize the expected utility of the portfolio's return.

c. **Black-Scholes Model and Delta Hedging**:

i. The model assumes the stock price follow the geometric Brownian motion: $dSt \coloneqq \mu Stdt + \sigma StdBt,\ 0 \leq t \leq T$

ii. By the assumption, we can calculate the call option price and by differentiate by the current stock price, we know the position we should hold. $D = \frac{\partial C}{\partial S_0} = N(d_1)$

$$C = SN(d_1) - Ke^{-rT}N(d_2)$$

$$d_1 = \frac{\ln(S/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}} \text{ and } d_2 = d_1 - \sigma\sqrt{T}$$

Required Inputs:

S = Current stock price
K = Option strike price
r = Risk-free interest rate
T = Time remaining until option expiration
σ = Volatility of the stock

d. **Reinforcement Learning Approach (DDPG)**:

DDPG is a is an actor-critic method suitable for continuous action spaces, combining value-based and policy-based methods.

3. Methodologies and Implementation

a. Data preparation

   i. Data source: The dataset consists of adjusted closed price of Apple Inc. (AAPL) stock obtain from Yahoo Finance.

   ii. Time Period: The data spans from 2012/1/1 to 2024/1/1.

   iii. Training and testing split: The first 80% of the data is used to training and the rest are for testing.

b. Environment setup

   i. For each episode, I the environment will randomly chooses a path as this time's interacting environment.

   ii. The payoff function is defined as $max(StrikePrice, StockPrice)$

c. Training and evaluation

   i. State space: (last time step stock price, volatility, current position, time to maturity)

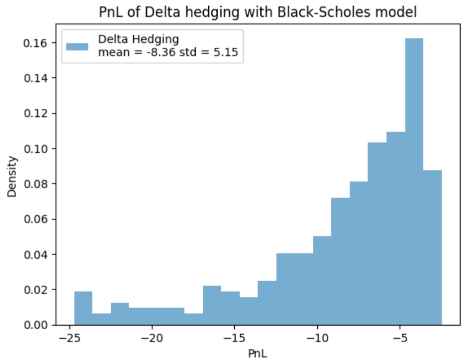
   ii. Action space: 0 to 1, which means the position of the stock.

   iii. Reward function: For each step, the reward is the value gain or loss after that action. And then put that in exponential utility function to get the reward.

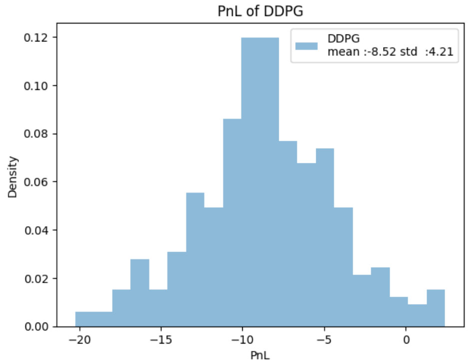
4. Result

DDPG reach roughly the same mean profit and loss and smaller standard deviation compared to Black-Scholes model delta hedging.

Black-Scholes delta hedging



DDPG



5. Conclusion and Future work

a. During the training process, the DDPG agent will sometimes outperform the Black-Scholes delta hedging, however, it does not converge after training for a long time.

b. The training dataset and testing dataset might not be in the same distribution, which might lead to the difficulty of training a RL agent.