Improving the Accuracy and Efficiency of the Longstaff-Schwartz Method for Pricing American Options through Machine Learning Techniques

Team 13: Jake Kemp, Jake Thakur, Mohammad Arif, and Uzochi Dimkpa

Problem Statement

- The problem we are tackling is pricing American options accurately. American options are a type of financial derivative that give the holder the right, but not the obligation, to exercise the option at any time before its expiration date. Accurately pricing American options is of great importance for both theoretical and practical reasons. The challenge in pricing these options lies in determining the optimal exercise strategy, given the underlying asset's price and other factors such as volatility and time to expiration. To solve this problem, we plan to incorporate machine learning techniques into the Longstaff-Schwartz method, which is a well-established approach for pricing American options using Monte Carlo simulation. By using machine learning, we hope to improve the accuracy of the pricing and reduce the vulnerability to misspecification of basis functions.

Introduction to Longstaff-Schwartz (LSM) Method

- The Longstaff-Schwartz method formulates the pricing problem as a two-step process. Firstly, it estimates the expected continuation value at each point in time, representing the expected future payout if the option were to be held. Secondly, it uses this information to make an exercise decision at each point in time, determining whether to exercise the option or not based on the expected future payout and the current option value. This dual formulation of the problem enables the Longstaff-Schwartz method to effectively capture the optimal exercise strategy for American options.

Pseudocode to illustrate the implementation of Longstaff-Schwartz (LSM) Method

The code is implementing the Longstaff-Schwartz method for pricing American options using the dual formulation. It does this in three steps:

- Simulate the underlying asset price paths using a random number generator and Monte Carlo simulation.
- Estimate the expected continuation value by backwards induction from the final time step to the initial time step.
- Calculate the option price by averaging the expected discounted payoffs at the initial time step.

```python
def longstaff_schwartz(S0, K, r, T, sigma, N, M, seed=None):
    # Step 1: Simulate the underlying asset price paths
    np.random.seed(seed)
    dt = T / N
    path = np.zeros((M, N + 1))
    path[:, 0] = S0
    for i in range(1, N + 1):
        z = np.random.normal(size=M)
        path[:, i] = path[:, i - 1] * np.exp((r - 0.5 * sigma**2) * dt + sigma * np.sqrt(dt) * z)

    # Step 2: Estimate the expected continuation value
    V = np.zeros((M, N + 1))
    V[:, -1] = np.maximum(K - path[:, -1], 0)
    for i in range(N - 1, 0, -1):
        X = path[:, i]
        Y = V[:, i + 1] * np.exp(-r * dt)
        beta = np.linalg.lstsq(X[:, np.newaxis], Y, rcond=None)[0]
        C = Y - beta * X
        exercise = K - path[:, i] > 0
        V[exercise, i] = K - path[:, i]
        V[~exercise, i] = beta * path[~exercise, i] + C.mean()

    # Step 3: Calculate the option price
    return np.average(V[:, 0] * np.exp(-r * T))
```

Motivation

- The accuracy and efficiency of pricing American options is of great importance to financial institutions and investors. Improved methods for valuing these derivatives can lead to better risk management and investment decisions. The potential of neural networks to improve the LSM method is the main motivation for this project.

Literature Surveys

Glasserman and Bin Yu's "Simulation for American Options: Regression Now or Regression Later?":

- This paper presents a comparison between two regression-based approaches for pricing American options, "regression now" and "regression later". The "regression now" approach estimates the option price using regression analysis at each time step, while the "regression later" approach first simulates the underlying asset and then uses regression analysis to estimate the option price. The authors conclude that the "regression later" approach is generally more efficient, as it avoids the need to perform regression analysis at each time step.

Longstaff and Schwartz's "Valuing American Options by Simulation: A Simple Least-Squares Approach":

- This paper proposes the Longstaff-Schwartz (LSM) method for pricing American options, which is a simulation-based method. The LSM method involves simulating the underlying asset and using least squares regression analysis to estimate the expected discounted payouts. The authors show that this method is efficient and can provide accurate results with a small number of simulations.

Jingying Lin and Caio Almeida's "American Option Pricing with Machine Learning: An Extension of the Longstaff-Schwartz Method":

- This paper proposes an extension of the Longstaff-Schwartz (LSM) method for pricing American options using machine learning. The authors show that using machine learning algorithms, such as artificial neural networks and support vector machines, can improve the accuracy of the LSM method by capturing non-linear relationships between the option price and the underlying asset. The authors also demonstrate that their method is computationally efficient and can handle large amounts of data.

Criticism of existing methods

1.) Glasserman and Bin Yu's "Simulation for American Options: Regression Now or Regression Later?":

- Limited scope: The comparison is only between two regression-based approaches and doesn't consider other methods.
- Limited applicability: The results are based on a limited set of option types and underlying assets and may not be generalizable to other scenarios.

2.) Longstaff and Schwartz's "Valuing American Options by Simulation: A Simple Least-Squares Approach":

- Assumptions: The method assumes that the underlying asset follows a known stochastic process and that the relationship between the option price and the underlying asset can be modeled using a linear regression, which may not always be accurate.
- Simplicity: The method is simple and efficient but may not be suitable for capturing complex relationships.

3.) Jingying Lin and Caio Almeida's "American Option Pricing with Machine Learning: An Extension of the Longstaff-Schwartz Method":

- Complexity: The use of machine learning algorithms may make the method more complex and harder to interpret for practitioners.
- Overfitting risk: There is a risk of overfitting the data when using machine learning algorithms, which can lead to poor results when the model is applied to new data.

Pros and Cons of existing methods

Approach	Pros	Cons
Glasserman and Bin Yu's "Simulation for American Options: Regression Now or Regression Later?"	- Comparison of two regression-based approaches, providing a useful reference for practitioners.	- Limited scope, only comparing two regression-based approaches. - Limited applicability, with results based on a limited set of option types and underlying assets.
Longstaff and Schwartz's "Valuing American Options by Simulation: A Simple Least-Squares Approach"	- Simple and efficient method for pricing American options.	- Assumes that the underlying asset follows a known stochastic process and that the relationship between the option price and the underlying asset can be modeled using a linear regression, which may not always be accurate.
Jingying Lin and Caio Almeida's "American Option Pricing with Machine Learning: An Extension of the Longstaff-Schwartz Method"	- Uses machine learning algorithms to improve the accuracy of the Longstaff-Schwartz method.	- More complex and harder to interpret than the standard Longstaff-Schwartz method. - Risk of overfitting the data when using machine learning algorithms, leading to poor results when the model is applied to new data.

Our approach compared to the existing literature

- Our approach to pricing American options using neural networks is related to previous methods such as Glasserman and Bin Yu's "Simulation for American Options: Regression Now or Regression Later?", Longstaff and Schwartz's "Valuing American Options by Simulation: A Simple Least-Squares Approach", and Jingying Lin and Caio Almeida's "American Option Pricing with Machine Learning: An Extension of the Longstaff-Schwartz Method". All of these methods utilize simulation and regression techniques to price American options but differ in their specific approach and techniques used.
- Glasserman and Bin Yu's method utilizes Monte Carlo simulation and compares the results of using regression now or regression later. Longstaff and Schwartz's method uses a least-squares approach in their Monte Carlo simulation. Lin and Almeida's method extends the Longstaff-Schwartz approach by incorporating machine learning techniques, including support vector regression and classification and regression trees, into the simulation process.
- Our approach is like these methods in that we also utilize simulation in the pricing process but differs in that we use neural networks instead of regression and classification techniques. The neural network is trained to predict the expected continuation value, which is then used in the exercise decision rule to determine whether to exercise the option or not. This allows for a more dynamic and adaptive model, potentially leading to improved pricing accuracy compared to previous methods.

Summary of Idea

- In our approach to pricing American options, we plan to incorporate deep learning techniques, specifically neural networks, into the Longstaff-Schwartz method. The objective of this method is to determine the expected continuation value at each time step and use it to make the exercise decision rule, deciding whether to exercise the option or not. The neural network will be trained to make these predictions, utilizing inputs such as underlying asset price, volatility, and time to expiration. The network will learn complex, non-linear relationships between the inputs and the expected continuation value, potentially leading to improved pricing accuracy compared to traditional methods. The use of neural networks in this manner represents a departure from previous studies which utilized regression and classification trees, as neural networks have shown promising results in a variety of financial applications. However, it is important to note that this approach may require more computational resources and data for training.

Questions we seek to answer

The questions we seek to answer by incorporating neural networks into the Longstaff-Schwartz (LSM) method for pricing American derivatives are:

- How does the integration of neural networks impact the accuracy of option pricing compared to the standard LSM method?
- Does the enhanced LSM method incorporating neural networks result in more robust and stable results compared to the standard LSM method?
- What is the computational time of both the standard LSM method and the enhanced LSM method incorporating neural networks and what is the trade-off between computational time and accuracy?
- How does the enhanced LSM method incorporating neural networks perform when tested on different types of underlying assets and market conditions compared to the standard LSM method?

What we will be able to learn?

By incorporating neural networks into the Longstaff-Schwartz (LSM) method for pricing American derivatives, we will learn:

- The effectiveness of using neural networks in modeling the underlying asset's price dynamics and the intrinsic value of the option compared to traditional regression methods used in the standard LSM method.
- How the integration of neural networks affects the accuracy, stability, and robustness of the option pricing model.
- The trade-off between computational time and accuracy and how the enhanced LSM method incorporating neural networks compares to the standard LSM method in terms of computational efficiency.
- The ability of the enhanced LSM method incorporating neural networks to handle diverse data and market conditions.

How is our approach different?

- Our approach differs from previous studies in several ways. Firstly, we incorporate deep learning techniques, specifically neural networks, into the Longstaff-Schwartz method for American option pricing. This is a departure from previous works which primarily utilized support vector regression and classification and regression trees. Secondly, our approach uses multiple input features such as underlying asset price, volatility, and time to expiration, allowing for a more comprehensive analysis of the data and improved pricing accuracy. Our method of training the neural network to predict the expected continuation value is also a unique aspect of our approach, as previous studies primarily used regression or classification trees to fit the data. Overall, our approach utilizes advanced machine learning techniques to enhance the pricing accuracy of the Longstaff-Schwartz method.

Plan

Task	Assigned To	Start Date	End Date
Literature Review	All Members	2/10/2023	2/20/2023
Data Collection	Jake K	2/20/2023	2/28/2023
Data Collection	Mohammad A	2/28/2023	2/28/2023
Data Collection	Uzochi D	2/20/2023	2/28/2023
Data Preprocessing	Jake K	2/28/2023	3/7/2023
Data Preprocessing	Jake T	2/28/2023	3/7/2023
Model Building	Jake K	3/7/2023	3/14/2023
Model Building	Jake T	3/7/2023	3/14/2023
Model Evaluation	Jake K	3/14/2023	3/17/2023
Model Evaluation	Uzochi D	3/14/2023	3/17/2023
Midterm Report	All Members	3/17/2023	3/17/2023
Model Improvement	All Members	3/17/2023	4/7/2023
Final Report	All Members	4/7/2023	5/5/2023

Citations

- Glasserman, P., & Yu, B. (2002). Simulation for american options: regression now or regression later?. Mathematical finance, 12(1), 1-22.
- Longstaff, F. A., & Schwartz, E. S. (2001). Valuing american options by simulation: A simple least-squares approach. The Review of Financial Studies, 14(1), 113-147.
- Lin, J., & Almeida, C. (2020). American option pricing with machine learning: An extension of the Longstaff-Schwartz method. Journal of Computational Finance, 23(4), 65-90.