Options Theory & Collecting/Analyzing Data

Kim Pham*[1, 2, 3]

[1] *Institute of Computing in Research*
[2] *Kiwan Hyun*
[2] *Mousumi Datta*

Abstract

This paper explores the application of options theory through a detailed analysis of historical options data using Python, with a focus on the Pandas library for data manipulation and analysis. The research involves gathering Flex Reports from the Options Clearing Corporation (OCC) and historical index data from Yahoo! Finance. These datasets are meticulously combined to create a historical record of options transactions and index movements, providing a foundation for further analysis.

Central to the study is the use of the Black-Scholes model, a fundamental tool in options pricing that allows for the estimation of theoretical option values, to back-solve for implied volatility. To refine this estimation, the paper employs Newton's method, a numerical technique used to solve for approximations for implied volatility with the help of the options Greek, Vega. This implied volatility is a crucial component for understanding market sentiment and is used to generate and analyze volatility smiles and volatility surfaces.

Volatility smiles and surfaces are graphical representations that illustrate how implied volatility varies with different strike prices and expiration dates. By plotting these metrics, the paper aims to reveal patterns and insights into market behavior and volatility expectations.

The findings from this analysis not only contribute to the academic understanding of options theory but also offer practical implications for trading strategies and risk management. By demonstrating how Python-based tools and historical data can be effectively utilized, the study highlights the value of quantitative analysis in the financial domain and provides a framework for further exploration in options pricing, volatility modeling, and possibly contributing to the applications of machine learning for predicting future stock prices and implied volatilities.

1 Introduction

Options theory, a critical component of financial economics, provides tools for pricing derivatives and managing financial risk. In order to get a good understanding on how options contracts work, let's first look at how stocks work. Stocks represent ownership in a company. When an individual buys a stock, they are purchasing a share of the company's equity, which entitles them to a portion of the company's profits and assets. Stocks are traded on stock exchanges, and their prices fluctuate based on market conditions. When stocks increase and the buyer decides to sell them, then the buyer would make a profit and vice versa for when the stocks decrease in price.

Options contracts, while related to stocks, function differently. A call option grants the holder the right, but not the obligation, to buy the underlying asset at a predetermined price before the expiration date. Conversely, a put option provides the right to sell the underlying asset under similar conditions. Both call and put options require the buyer to pay a premium to acquire these rights.

Within the options market, participants can hold either long or short positions. A long position signifies the acquisition of the options contract, whereas a short position involves writing or selling the contract. Long position holders are obliged to fulfill the

*kim.kp.pham@gmail.com

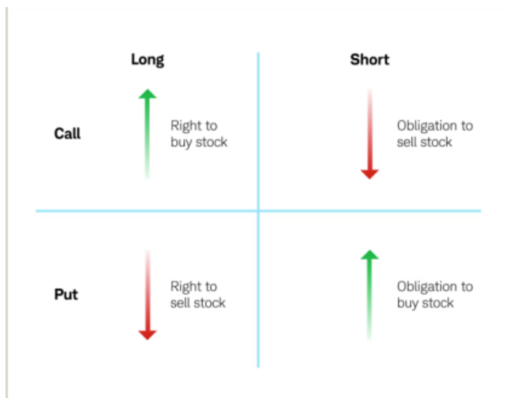


Figure 1: Longs, shorts, puts, and calls in different combinations

contract's terms if the short position holder decides to exercise their option. The financial outcomes for each position vary depending on the movements in the underlying asset's price.

Options contracts are defined by their strike price, which is the price at which the underlying asset can be bought or sold if the option is exercised. For instance, a long position in a call option with a strike price of $145 allows the holder to purchase the underlying asset at this lower price, despite its current market value being $200, thereby realizing a profit.

To assess the fairness of an options contract price and assist in contract valuation, various financial models are employed. Among these, the Black-Scholes model is particularly significant. It offers a theoretical framework for pricing European-style options, assuming factors such as constant volatility and efficient markets. However, this model requires an estimate of the underlying asset's volatility, which is not directly observable. This is where implied volatility becomes crucial.

Implied volatility is the volatility value that makes the theoretical price from the Black-Scholes model match the observed market price of the option. Estimating implied volatility requires solving for this value, and Newton's method, a powerful numerical technique, is commonly used for this purpose. Newton's method iteratively refines an initial estimate of volatility until the theoretical price aligns with the market price.

Furthermore, once implied volatilities are determined, they can be used to visualize how volatility varies across different strike prices and expiration dates. Graphical tools such as volatility smiles and surfaces provide insights into market expectations. A volatility smile plots implied volatility against strike prices for a fixed expiration date, often showing a "smile" shape where volatility is higher for options that are deep in-the-money or out-of-the-money. The volatility surface extends this analysis by mapping implied volatility against both strike prices and expiration dates, offering a comprehensive 3D view of volatility trends across various dimensions.

2 Data Collection

To effectively estimate implied volatility using the Black-Scholes model, it is essential to systematically collect and integrate relevant data. For this analysis, data was sourced from the Options Clearing Corporation (OCC) and Yahoo Finance, which provided comprehensive information on options contracts and historical prices of the underlying asset.

The primary data source was the OCC's open interest flex reports, which offered detailed information about options contracts. This dataset included the mark price of the options, which represents the most recent trading price and reflects the current market value of the options. Additionally, the strike price of each options contract was recorded, indicating the price at which the underlying asset can be bought or sold under the contract terms. The contract date, when the options were created or first traded, and the expiration date of the options, marking the deadline by which the option must be exercised, were also included in the reports.

To calculate the time to expiration, which is crucial for the Black-Scholes model, the number of days remaining until the expiration date was determined. This time span was converted into years by dividing by 365, providing a standardized input for the model.

Historical price data for the underlying asset was obtained from Yahoo Finance. This dataset included the historical prices of the asset on specific dates, allowing for a contextual understanding of the asset's price movements over time. By aligning these historical prices with the dates of the options data, a coherent dataset was created that facilitated accurate analysis.

3 Black-Scholes Model

The Black-Scholes model provides a formula for pricing European call and put options. The model assumes that the underlying asset follows a geometric Brownian motion with constant volatility and that there are no arbitrage opportunities. The formula for a European call/put option is given by Figure 2 where C is the price of the call option, P is the price of the put option, S is the current price of the underlying asset, r is the risk-free interest rate, T is the time to maturity, and N is the combined distribution function of the normal distribution. These variables will be described in more detail in the following section.

Here, sigma represents the volatility of the underlying asset, which is not directly observable but derived by combining the Black-Scholes model with Newton's method once all of the other variables are known.

$$C(S_t, t) = S_t N(d_1) - Ke^{-r(T-t)} N(d_2)$$
$$P(S_t, t) = Ke^{-r(T-t)} N(-d_2) - S_t N(-d1)$$
$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

Figure 2: Black-Scholes Model

4 Newton's Method in Calculating Implied Volatility

Newton's method is employed to solve for implied volatility by iteratively refining an estimate to minimize the difference between the theoretical price of the option, as calculated by the Black-Scholes model, and the actual market price. The theoretical price of an option, determined using the Black-Scholes formula, depends on several parameters including the strike price, time to expiration, risk-free interest rate, and volatility. Since volatility is not directly observable, it must be inferred by solving the inverse problem where the theoretical price matches the observed market price.

4.1 Overview of Newton's Method

Newton's method, also known as the Newton-Raphson method, is an iterative procedure used to find approximations to the roots of a real-valued function. The method operates on the principle of linear approximation, utilizing both the function values and its derivatives to improve the accuracy of the estimate.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Figure 3: Newton's Method

Given a function and its derivative $f\prime(x)$, Newton's method refines an initial guess x_0 for the root using the following iterative formula:

As seen in figure 3, x_n represents the current estimate of the root, and $x_n + 1$ is the updated estimate. The method uses the tangent line at x_n to approximate where the function intersects the x-axis, thus providing a new estimate.

In the context of estimating implied volatility, the function $f(\sigma)$ is defined as the difference between the theoretical price $C_{\text{BS}}(\sigma)$, computed using the Black-Scholes model, and the observed market price C_{market}

5 Plotting Volatility Smiles and Surfaces

5.1 Volatility Smiles

In our analysis, we visualized volatility smiles to understand how implied volatility varies with different strike prices for a given expiration date. To achieve this, we utilized a pandas DataFrame and Matplotlib.

The dataset used for plotting comprised several columns, including expiration dates, time to expira-

```python
# Filters for rows for matching expiration/time-to-expiration/flag
df = df[df['EXPIRATION']==expiration]
df = df[df['TIME-TO-EXPIRATION']==time]
df = df[df['P/C']==flag]

# Extract columns
x = df['STRIKE-PRICE']
y = df['IV']

# Create scatter plot
plt.figure(figsize=(20, 12))
plt.scatter(x, y, marker='o')
plt.plot(x, y, marker='o')

# Formatting
plt.title(plot_title)
plt.xlabel('STRIKE-PRICE')
plt.ylabel('IV')

plt.xticks(rotation=45)
plt.gca().xaxis.set_major_locator(plt.MaxNLocator(15))  # Adjust the number of ticks shown

plt.tight_layout()
plt.grid(True)
plt.show()
```

Figure 4: Constructing Volatility Smile

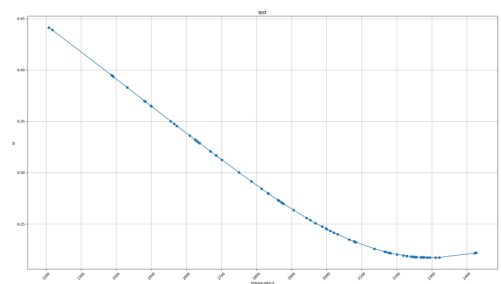


Figure 5: Volatility Smile Example

tion, and the type of option contract. The primary focus was on the implied volatility of options, which was calculated using the Black-Scholes model and refined through Newton's method. For each expiration date, we extracted the relevant strike prices and their corresponding implied volatilities.

Using Matplotlib, we generated scatter plots where the x-axis represented strike prices and the y-axis depicted the implied volatilities. This visualization allows us to observe the volatility smile—a characteristic curve showing how implied volatility tends to increase for both deep in-the-money and out-of-the-money options compared to at-the-money options.

The resulting plot illustrated the typical "smile" pattern, providing insights into market expectations and hedging behaviors. Such visualizations are crucial for understanding volatility dynamics and enhancing options pricing strategies.

5.2 Volatility Surfaces

To gain a comprehensive understanding of how implied volatility varies with different strike prices and time to expiration, we visualized the volatility surface using a three-dimensional plot. This method provided a more nuanced view of volatility dynamics compared to the traditional volatility smile.

The analysis was based on a dataset containing strike prices, time to expiration, and the corresponding implied volatilities. By plotting these variables, we constructed a three-dimensional surface where the x-axis represented the time to expiration, the y-axis denoted the strike price, and the z-axis depicted the implied volatility.

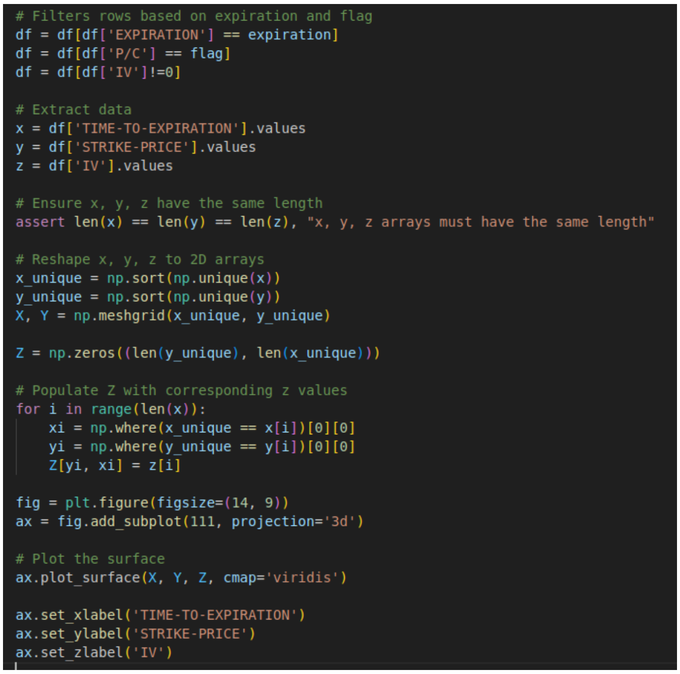
The 3D plot was generated using Matplotlib, which allowed for an intuitive representation of how implied volatility evolves across different dimensions. This visualization method enabled us to observe the intricate patterns of volatility as it changes with both strike prices and expiration dates.

The resulting surface plot illustrated how implied volatility generally exhibits complex, non-linear behavior depending on the strike price and the time remaining until expiration. Such visualizations are instrumental in capturing the broader context of market expectations and pricing strategies, providing a deeper insight into the factors influencing options pricing.

By employing this three-dimensional representation, we were able to better understand the interaction between volatility, strike prices, and expiration timelines, thereby enhancing our analysis of market conditions and risk management strategies.

6 Conclusion

This study examines options theory through the analysis of historical options data using Python. Data sourced from the Options Clearing Corporation (OCC) and Yahoo Finance facilitated the evaluation of options pricing and implied volatility. The Black-Scholes model was applied to estimate theoretical option prices, while Newton's method was employed to derive implied volatility. This iterative process aligned theoretical prices with observed mar-



```python
# Filters rows based on expiration and flag
df = df[df['EXPIRATION'] == expiration]
df = df[df['P/C'] == flag]
df = df[df['IV']!=0]

# Extract data
x = df['TIME-TO-EXPIRATION'].values
y = df['STRIKE-PRICE'].values
z = df['IV'].values

# Ensure x, y, z have the same length
assert len(x) == len(y) == len(z), "x, y, z arrays must have the same length"

# Reshape x, y, z to 2D arrays
x_unique = np.sort(np.unique(x))
y_unique = np.sort(np.unique(y))
X, Y = np.meshgrid(x_unique, y_unique)

Z = np.zeros((len(y_unique), len(x_unique)))

# Populate Z with corresponding z values
for i in range(len(x)):
    xi = np.where(x_unique == x[i])[0][0]
    yi = np.where(y_unique == y[i])[0][0]
    Z[yi, xi] = z[i]

fig = plt.figure(figsize=(14, 9))
ax = fig.add_subplot(111, projection='3d')

# Plot the surface
ax.plot_surface(X, Y, Z, cmap='viridis')

ax.set_xlabel('TIME-TO-EXPIRATION')
ax.set_ylabel('STRIKE-PRICE')
ax.set_zlabel('IV')
```

Figure 6: Constructing Volatility Surface

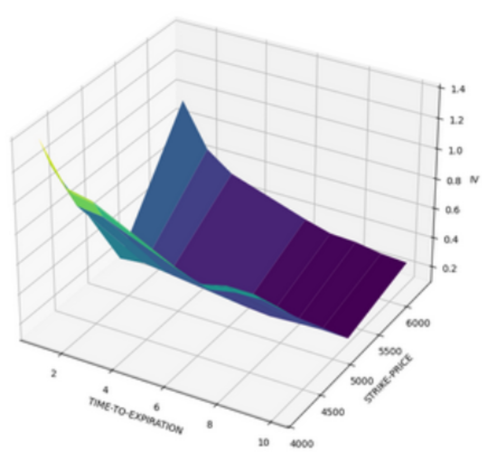


Figure 7: Enter Caption

ket prices, providing accurate volatility estimates.

Volatility smiles and surfaces were constructed to visualize variations in implied volatility. The volatility smile revealed that implied volatility tends to increase for options that are either deep in-the-money or out-of-the-money compared to at-the-money options. The 3D volatility surface offered insights into how implied volatility varies with strike prices and time to expiration, highlighting complex market patterns.

The findings emphasize the value of quantitative analysis in financial markets and showcase the effective application of Python tools for options pricing and volatility assessment. The research lays a foundation for further exploration and the potential use of advanced techniques, such as machine learning, for enhanced forecasting and risk management.

References

[1] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637-654. https://doi.org/10.1086/260062

[2] Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7*(3), 229-263. https://doi.org/10.1016/0304-405X(79)90015-1

[3] Hull, J. C. (2017). *Options, futures, and other derivatives* (10th ed.). Pearson.

[4] Jarrow, R. A., & Rudd, A. (1983). *Option pricing: A simplified approach*. Harper & Row.

[5] Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science, 4*(1), 141-183. https://doi.org/10.2307/3003143

[6] Options Clearing Corporation (OCC). (n.d.). *Flex reports*. Retrieved from https://www.theocc.com/

[7] Yahoo Finance. (n.d.). *Historical stock data*. Retrieved from https://finance.yahoo.com/

[8] Wilmott, P. (2006). *Paul Wilmott introduces quantitative finance* (2nd ed.). Wiley.