Applications of Gaussian Process Regression to Portfolio Pricing

Julian Chitiva*

December 22, 2022

1 Introduction

Financial Markets are an important part of our society. Thanks to them we are able to trade objects around the world and allow the economic integration and development of countries. To have a measure of the importance of these trades we can observe the volume[1] of transaction made by the NYSE and NASDAQ in U.S.A was 3,847,676,312 and 4,566,906,059, respectively. Given this relevance, it is important to have appropriate methodologies for pricing these instruments. The results from these models could in decisions about investment and hedging strategies. It is well known that, some of the most important financial instruments to hedge against risk are the Options. Is because of this, that during this work we will restrict our analysis to these family of derivatives.

The financial options are usually priced using the Black-Scholes model, which created by Fischer Black and Myron Scholes during 1960's (see Section 2 for the details). This model gives the theoretical price of an European option[2] based on observations of some characteristics of the underlying asset and parameters of the option. Nevertheless, in reality, investors could only partially observe these information which raises a concern for correct pricing of this instruments. In this context Gaussian Process Regression (GPR) plays a role because it offers a solution to the pricing problem. In particular, the GPR model helps with the selection of the pricing model by letting the data decide on the complexity of the function. The scope of this work is to use the GPR Model to option price data. To do so, we first need to get some data points which are generated by the Black-Scholes Formula in order to train the GPR model, then we use this to predict the price of the financial instruments. Up to this moment, we have only talked about some generic European Options, now we can define the ones that will be in the scope of this work. We will start with *plain vanilla* world of derivatives, which is have most basic financial instruments. This world is composed by *call* and *put* options.

A *call* option is a financial instrument that grants its owner the right to buy a given amount (the notional) of an asset (the subjacent or underlying asset) for an unitary given price (the strike) in a given future date (expiration date). These instruments are traded in organized markets (OTC) via bilateral contracts. In these contracts there are two sides: the *short* side, which is the one that sells the option or give the rights to the *long* side, which is the one that buys the option. The important difference between the options and other financial instruments like the futures or forwards is that the owner has the right to

*HEC Paris julian.chitiva@hec.edu - ENSAE julian.chitiva@ensae.fr

[1]See Wall Street Journal

[2]*European* options could only be exercised only at the expiration date. *American* options, could be exercised at any time before the expiration date.

decide to buy the underlying asset. If the owner of the option decides to buy the asset according to the contract conditions we say that the owner exercise the option. One can note that the owner of the option will exercise the option at the expiration date (T) if the price of the underlying asset ($S(T)$) is higher than the strike (K). Otherwise, it will be more profitable to buy it directly from the market. Therefore, the owner of the *call* option will receive $\max(0, S(T) - K)$, as shown in Figure B5a. An individual may be motivated to buy a *call* option if she wants to participate on increases in prices, limit losses or leverage profits.

On the other side, the *put* options will follow the same definitions and satisfy every condition from the *call* option with the only difference that the owner has the right to sell the notional of the underlying asset at the strike price in the expiration date. Therefore, the owner of the *put* option will receive $\max(0, K - S(T))$, as shown in Figure B5b. An individual may be motivated to buy a *put* option if she wants to participate on decreases in prices, limit losses or leverage profits.

Even though these *plain vanilla* derivatives are the basis of the financial markets, this work extends the analysis to other more exotic derivatives. In particular, to *Call Spread* or *Bull Spread*, *Put Spread* or *Bear Spread*, *Straddle*, *Strangle* and*Butterfly Spread*. These derivatives are traded on the market and are compositions or aggregation of multiple *put* or *call* options. It is important to notice that each one of them constitute more complex hedging strategies to limit risk in certain regions of the price of the underlying asset.

First, the *Bull Spread* is when an investor simultaneously buys a *call* at a specific strike price K_1 while also selling a *call* at a higher strike price K_2. It is important to notice that both *call* options will have the same expiration date and underlying asset. An investor may be interested in buying this derivative strategy when she expects a moderate rise in the price of the asset. Using this strategy, she is able to limit their upside on the trade while also reducing the amount spent when buying the options as shown in Figure B5c.

Second, the *Bear Spread* is when an investor simultaneously purchase a *put* option at a specific strike K_1 price and also sells a *put* at a lower strike price K_2. It is important to notice that both *put* options will have the same expiration date and underlying asset.An investor may be interested in buying this derivative strategy when she expects a moderate fall in the price of the asset. The strategy offers both limited losses and limited gains as shown in Figure B5d.

Third, the *Straddle* is when an investor simultaneously purchase a *call* and *put* option on the same underlying asset with the same strike price and expiration date. The investor may use this strategy when she believes the price of the underlying asset will move significantly out of a specific range, but they are unsure about the direction. Theoretically, this strategy allows her to have the opportunity for unlimited gains with limited losses as shown in Figure B5e.

Fourth, the *Strangle* is when an investor simultaneously purchase a the investor purchases a call and a put option with a different strike price K_1 and K_2, respectively, on the same underlying asset with the same expiration date. The investor may uses this strategy when she believes the underlying asset's price will experience a very large movement but is unsure of which direction the move will take. Because of the different in strikes with the *Straddles*, the *Strangle* are usually cheaper and therefore with lower losses as shown in Figure B5f.

Finally, the *Butterfly Spread* is when an investor simultaneously purchase a one in-the-money *call* option at a lower strike price K_1, while also selling two at-the-money *call* options and buying one out-of-the-money *call* option. The investor may buy a *Butterfly Spread* when they think the stock will not move much before expiration because the maximum gain is when the price of the subjacent remains unchanged, as shown in Figure B5g.

2 Methodology

To tackle the problem of portfolio pricing we need first to define the basic theoretical framework to price options. The Black-Scholes model (Black and Scholes, 1973) is a mathematical model for the dynamics of a financial market that contains financial derivatives.

From this model one can deduce Equation 1 known as the Black-Scholes formula, which gives a theoretical estimate of the price $C(S_t, t)$ of the European *call* options at time t. In order to price European *put* options one could use the put-call parity (Equation 2) to deduce the price $P(S_t, t)$ according to the Black-Scholes model, as shown in Equation 3. The Black-Scholes formula shows that the option has a unique price given the risk of the security and its expected return. The intuition behind this model is the idea of hedging the option by buying and selling the underlying asset in a specific way to eliminate risk. This type of hedging is called *continuous revised delta hedging* and is the basis of more complicated hedging strategies like the ones used by investment banks and hedge funds.

Finally, investors also focus on some other metrics called Greeks which gives a notion of the change in price with respect to the change of some parameters of the derivative. This work will focus on Delta and Vega, which are ones of the most used by the Financial Institutions. First, Delta (δ) is the change of the price of the option (or a derivative) with respect to the price of the underlying asset as shown in Equation 4. For example, for the European *call* option, the Equation 5 states the δ from the Black-Scholes Model. Next, given the importance of the implied volatility for the correct pricing of these derivatives, Vega (ν) which is the change in price with respect to changes in the implied volatility, as shown in Equation 6. It is important to note that, one key assumption from the Black-Scholes Model is that σ was constant, so how could we justify ν? In practice, the volatility σ that is used by financial institutions to trade and price this instruments change adapting to expectations of future volatility.

Since Black-Scholes is the theoretical standard in pricing of options, we will use this model to generate data for prices to feed the Gaussian Process Regression (GPR) Model. From this point, the idea is to fit a GPR Model to price the financial instruments of interest. The motivation to use this regression model is that some investor wants to price an option observes partial information about the value of the option and the price of the underlying asset. From this scenario the investor needs to predict the value of the financial instrument possibly without knowing some additional information like the implied volatility, for example. She could fit a GPR model to the data to interpolate the value of the option and take more informed decisions. In this report, we will not enter in the details of the GPR Model since we cover them in class. Nevertheless, it is important to mention that the we will use the Radial basis function kernel for the calculations, which has a natural interpretation of similarity.

$$C(S_t, t) = \mathcal{N}(d_1)S_t - \mathcal{N}(d_2)Ke^{-r(T-t)} \tag{1}$$

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right) \times (T-t)\right]$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{z^2}{2}}\mathrm{d}z$$

S_t : Price of underlying asset at time t.

σ : standard deviation of stock returns.

r : risk-free interest rate.

K : Strike of the option.

$$C(t) - P(t) = S_t - Ke^{-r(T-t)} \tag{2}$$

$C(t)$: Price of *call* option at time t.

$P(t)$: Price of *put* option at time t.

$$\begin{aligned} P(S_t, t) &= Ke^{-r(T-t)} - S_t + C(S_t, t) \\ &= \mathcal{N}(-d_2)Ke^{-r(T-t)} - \mathcal{N}(-d_1)S_t \end{aligned} \tag{3}$$

$$\delta = \frac{\partial \text{price}}{\partial S} \tag{4}$$

$$\delta(call) = \frac{\partial C(S_t, t)}{\partial S_t} = \mathcal{N}(d_1) \tag{5}$$

$$\nu = \frac{\partial \text{price}}{\partial \sigma} \tag{6}$$

3 Code Execution

The code runs in the "JupyterLab + LaTeX" application of the Nuvolos web app. It is composed by the following items:

- `BlackScholes.py` python file that contains the functions for the Black-Scholes formula for pricing options.
- `aux_functions.py` python file that contains the functions to define classes for the different financial instruments and portfolio. For a generic Derivative, it contains the functions to fit the Gaussian Process Regression Model and compare this results with the Black-Scholes Model. It also contains the functions to price portfolios and the functions to analyze the results.
- `portfolio.txt` is a text file that contains information about the portfolio. This file is the input for more complex portfolios and can be modified to obtain different results.
- `portfolio_class.txt` is a text file that contains information about the portfolio from the class. This file is the input for more complex portfolios and can be modified to obtain different results. This portfolio is composed by long two *call* options and short one *put*.
- `FinalProject_JulianCHITIVA.ipynb` is the Jupyter Notebook that executes the code for the different financial instruments and an example of a portfolio. This is the main Notebook that contains the executed code and the examples.

4 Results

In this section, we will analyze the results for the *call* and *put* options and the portfolio from the class. The analysis for the other instruments will follow from this since they are compositions of the *plain vanilla* options and they can be found in the notebook `FinalProject_JulianCHITIVA.ipynb`.

First, we observe that for the price of the options it is not necessary to have a large training sample. Nevertheless, it is always recommended to have bigger training sample because of the effect of the number of options that a portfolio can hold, as we can observe in Figure A1. For example an error of US$3 in the price of the option could become into millions of dollars depending on the number of options that the investor has.

The case of the Greeks of these options is a bit different. We observe that a small training sample size leads to bad approximations of them. This is probably a result of some propagation of the error of the estimation of the value of the option. We can observe this Figure A2 and A3.

Finally, as an example for the functionality of the portfolio we can observe the results for the class' portfolio in Figure A4.

5 Conclusions and future work

During this project we extend the functionalities of the Gaussian Process Regression Model from pricing *put* and *call* European options to price *Bull Spread*, *Bear Spread*, *Straddle*, *Strangle* and *Butterfly Spread*, and in general more complex derivatives portfolios. In Section 4 we observe the Valuations, Delta and Vega of the different financial instruments. In addition, we observe the aggregation of them into a portfolio. It is important to mention that the size training sample has a key role in the correct estimation of this values. This parameter has a different magnitude effect which depends on the variable that we want to predict. On one side, the effect on the price is small and we observe convergence to the theoretical value even with small training samples. On the other side, the effect on second order variables like δ or ν is big because of propagation or amplification effect of the error from previous steps like the pricing.

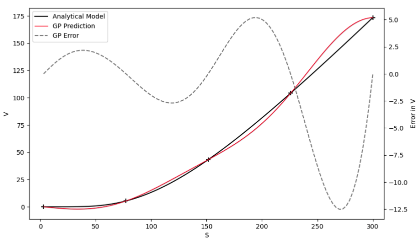
Finally, there are two possible additional extensions to the analysis made in this work. First, is to include some dynamic valuation. This is a natural extension because investors are interested in knowing the value of their portfolio across time to calculate their P&L and adjust their investment decisions. This implementation would need simulations of the price of the underlying asset which could be done by implementing a Geometric Brownian Motion or an econometric model and adjusting the mean of the simulation to some financial instruments like Futures to capture the information that the market has. Second, given the importance of the risk management for the investors this model could be extended to find optimal hedging strategies given a portfolio of commodities. One could develop some hedging strategies using the financial instruments considered in this project to answer questions like, given my exposition to the risk of commodity x it is better to buy or sell a *call* option or a *bull spread*? And with this information build the optimal portfolio of financial instruments to hedge the risk of a portfolio of commodities.

References

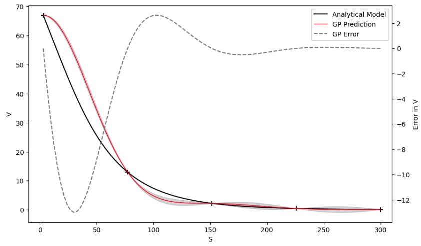
Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654.

Appendix

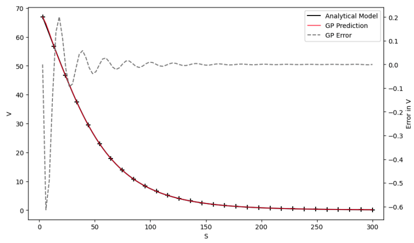
A Gaussian Process Regression Results



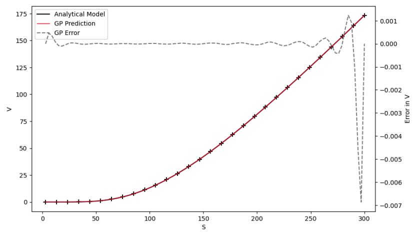
(a) Call Option. training_number=5



(b) Put Option. training_number=5

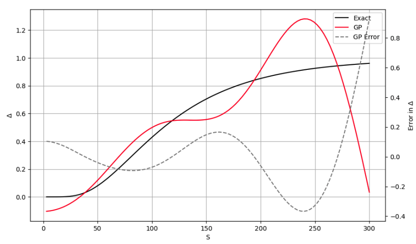


(c) Call Option. training_number=30

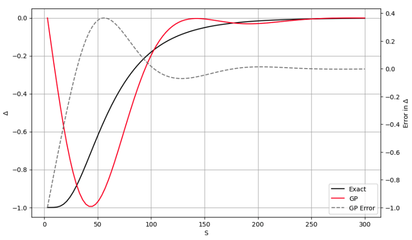


(d) Put Option. training_number=30

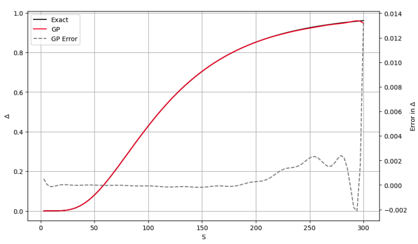
Figure A1: Value of Options



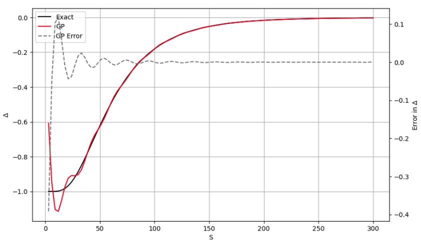
(a) Call Option. training_number=5



(b) Put Option. training_number=5

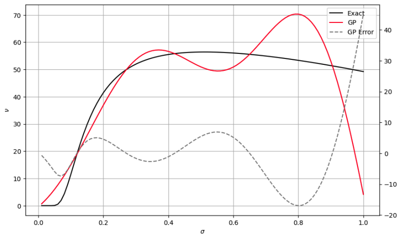


(c) Call Option. training_number=30

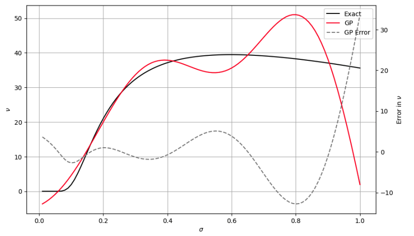


(d) Put Option. training_number=30

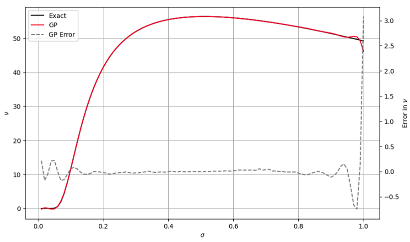
Figure A2: Delta of Options



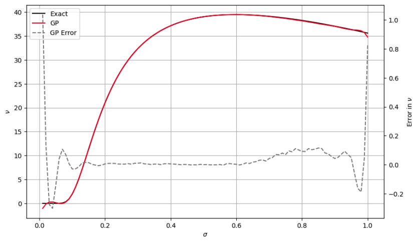
(a) Call Option. training_number=5



(b) Put Option. training_number=5

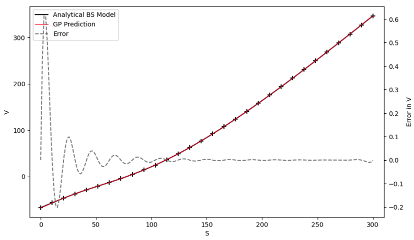


(c) Call Option. training_number=30

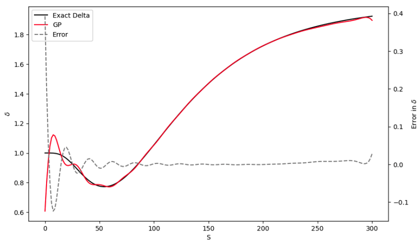


(d) Put Option. training_number=30

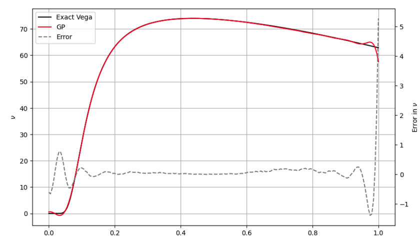
Figure A3: Vega of Options



(a) Value



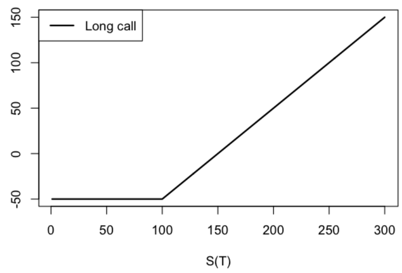
(b) Delta



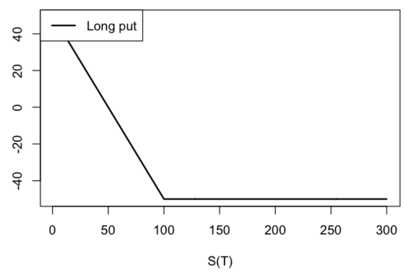
(c) Vega

Figure A4: Class' Portfolio.

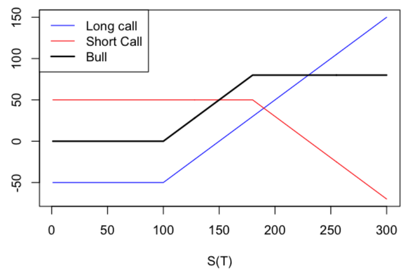
B Profits and Looses (P&L) plots



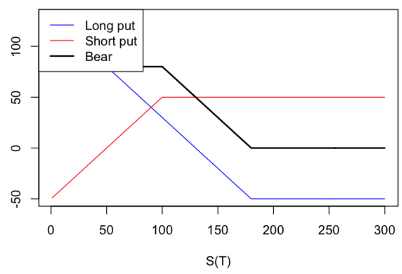
(a) Call Option



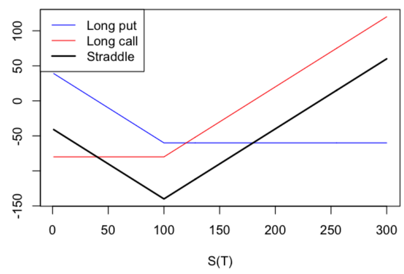
(b) Put Option



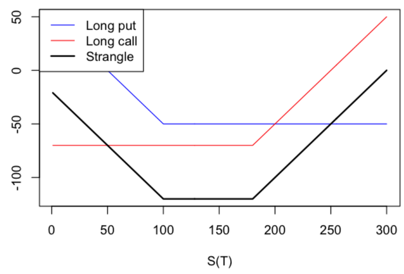
(c) Bull Spread



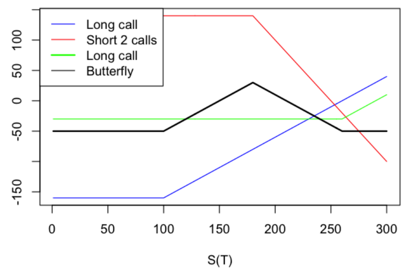
(d) Bear Spread



(e) Straddle



(f) Strangle



(g) Butterfly Spread

Figure B5: P&L of derivatives based on options strategies.