MATH-GA - 2707
Ziming Wang zw1754
Yuqing Wang yw3637
Instructor: Farshid Maghami ASL
Instructor: Rob Reider
Dec 21, 2019

Time Series and Statistical Arbitrage Project

Volatility Estimation and Option Hedging Trading Strategy

Contents

1. Introduction

1.1. Purpose

There are studies indicates that GARCH type model does not predict as well as precise as historical volatilities, but there are few studies about how GARCH type model performs in trading. This project explores whether GARCH type model could provide additional information as an advantage to generate a stable profit.

1.2. Volatility Estimation

We explore the feature of data: continuity, stationary, autocorrelation, mean reversion and kurtosis of data to make sure it is reasonable to build GARCH type models on daily returns. We build four models: ARCH, GARCH, T-GARCH and I-GARCH models for volatility estimation. Max likelihood estimation is used to estimate parameters and Schwartz Bayesian Criteria are used for model assessment and selection.

1.3. Backtesting and Explanation

In this part we collected 5 call options corresponding with the underlying stocks which we used to build the models. We then calculated the implied volatility of those options. By comparing the implied volatility of those options with the "true volatility" by our model, we can then form a trading strategy which sell calls when the implied volatility is higher and buy calls when the implied volatility is lower than our predicted volatility. To make our strategy neutral with stock price's movement, we let our portfolio delta neutral when we enter a trade. Our result indicates that this trade strategy can generate positive return.

2. Data

2.1. Data Description

Options with dividends are always difficult to tackle in pricing problems. In order to get rid of dividends related problems. Five non-dividend stocks: Amazon.Com Inc. (AMZN), Facebook Inc. (FB), Electronic Arts Inc. (EA), Netflix Inc. (NFLX) and Adobe Systems Inc. (ADBE) are selected. The data sets are their historical stock prices and their almost at the money call options. Stock price data starts from 2016-01-04 and end at 2019-12-20: there are 1000 data for each stock.

The following chart shows fundamental information of stocks at the money option price data:

	AMZN	**FB**	**EA**	**NFLX**	**ADBE**
Start Date	2019-11-18	2019-11-18	2019-10-30	2019-10-30	2019-10-31
Start Time	10:10:00	09:30:00	09:30:00	09:30:00	09:30:00
End Date	2019-12-16	2019-12-16	2019-12-16	2019-12-16	2019-12-16
End Time	11:30:00	12:00:00	12:10:00	12:10:00	12:30:00
Total	290	540	494	917	274

Table 1 Data description

It is hard to find option data for long period: they are not supplied from Bloomberg. So high frequency option data are used for backtesting in this project.

2.2. Data Exploration

In this section, data characteristics are detected for model selection. And only FB stock data before 2019-11-18 are used to avoid overfitting or cheating.

2.2.1. Frequency

Based on data, there are two main choices for the frequency for the time series analysis: either 10 minutes or daily. The following graph compares the plot of these two frequencies:

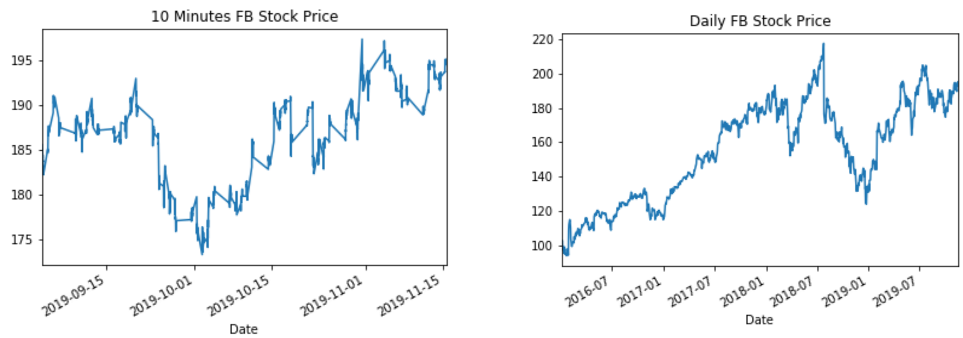

Figure 1 10 Minutes FB stock price and daily FB stock price

The left graph shows 10-minutes FB stock price. The next morning's price is right behind this afternoon price. It is obvious that the left graph has more jumps than the right one: one could trade only one third of the day: so, there is much more uncertainty happen during the close-market hours. For daily stock price, it covers almost 5/7 of the whole time.

2.2.2.Stationary

Instructed by Professor ASL, it would be better to build models on returns other than simply stock price. And it is obvious to us that the price is trendy. The following graph shows the partial autocorrelation of the daily log return of FB stock:

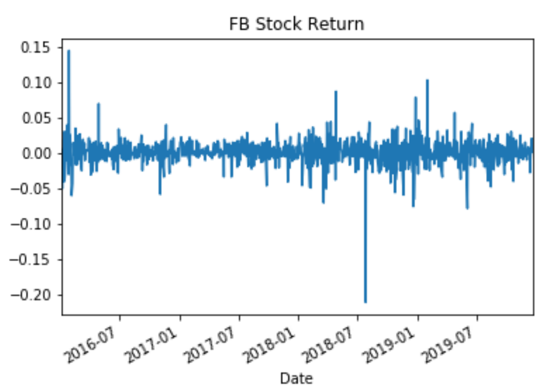

Figure 2 FB stock return

By calculating the return, the time series model is detrended.

2.2.3.Autocorrelation

For GARCH type models, the time series should itself has zero autocorrelation, but the square of the return should have positive autocorrelations. Here are the graphs:

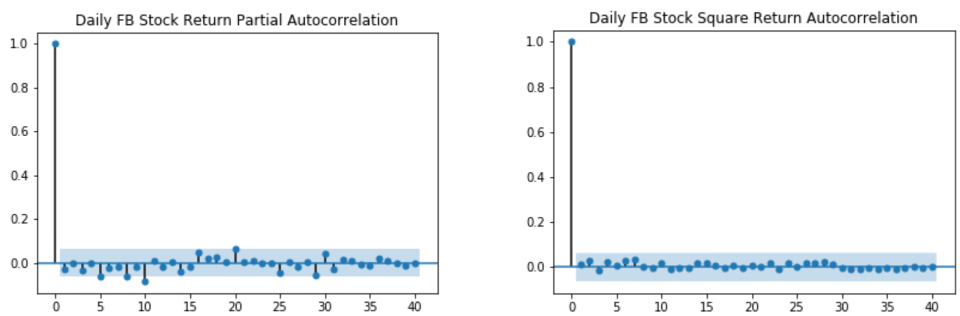

Figure 3 FB stock return autocorrelation and partial auto correlation

The left graph shows autocorrelation of the returns. The autocorrelation is not significant as GARCH type model requires. The autocorrelation of square return is not significant as well, but at least it goes positive of the order one and two.

2.2.4.Mean Reverting and Extra Kurtosis

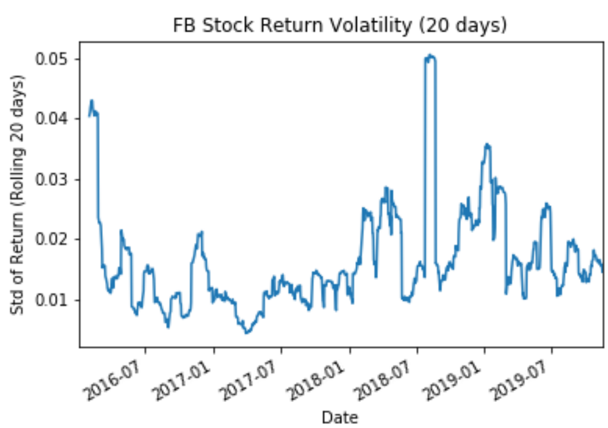

Figure 4 20 days rolling volatility

This is the graph of rolling 20 days standard deviation or return. The mean reverting tendency exists, and the kurtosis is 23.29, which is significantly larger than 3. So, GARCH type model would be a good fit.

3. Volatility Estimation

3.1. Models

We built four models: ARCH, GARCH, T-GARCH and I-GARCH models for volatility estimation: the aim is to maximize the likely likelihood:

$$L(parameters \mid r_1, r_2, \ldots, r_T) = \prod_{i=1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} exp(\frac{-(r_i - \mu)^2)}{2\sigma_i^2})$$

$r_i$is the log return of time i and $\mu$is the mean of return which is one of the parameters needs estimating.

For different models, σ_i^2 are estimated in different ways while σ_0^2 are initialized as the long-term standard deviation of log return:

MODEL	**METHOD**
ARCH	$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$
GARCH	$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$
T-GARCH	$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \gamma_1 S_{t-1} a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$
I-GARCH	$\sigma_t^2 = \alpha_0 + (1 - \beta_1) a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$

Table 2 Different model for volatility

where α_0, α_1,γ_1 and $\beta_1$are the parameters to estimate;
Python third party package scipy.optimize.minimize with nelder-mead optimization method is used for solving the optimization problem.

3.2. Parameter Estimation Result - FB as an Example

The following table shows the parameters for models of FB:

	$\boldsymbol{\mu}$	$\boldsymbol{\alpha_0}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\beta_1}$
ARCH	6.12e-04	2.62e-04	3.41e-01	-	-
GARCH	2.38e-04	8.63e-05	3.54e-01	-	4.93e-01
T-GARCH	5.19-05	1.01e-04	2.68-01	1.55e-01	4.55e-01
I-GARCH	7.77e-05	6.15e-05	-	-	5.35e-01

Table 3 volatility model parameter estimation

The optimization results are shown in the following chart:

	LIKELIHOOD	N PARAMETERS	SBC
ARCH	2526.49	3	12.81
GARCH	2545.81	4	19.69
T-GARCH	2546.55	5	26.57
I-GARCH	2543.01	3	12.80

Table 4 Time series model result

Even without Schwartz Bayesian Criteria, it is not hard to guess that I-GARCH would perform best. For one thing, with only 3 parameters it behaves as well as GARCH even T-GARCH. For another, estimated μ is 7.77e-05, which is closest to the mean of real real stock return.

4. Model Assessment

As mentioned in the introduction, we collected 5 call options corresponding to the stocks. They are, call with strike 1757.5 for AMZN, call with strike 197.5 for FB, call with strike 150 for EA, call with strike 305 for NFLX and call with strike 325 for ADBE. All of them expire on 12/20/2019 16:00. The frequency of option price is 10 minutes, but due to the lack of liquidation, we cannot guarantee that we can get a price every 10 minutes.

The first step we do in this part is to calculate the implied volatility, as well as the delta of the options. We plot the implied volatility of those options below and we found that 2 of them (AMZN, ADBE) did not have enough liquidity to trade. So, we did not use them in our trading strategy.

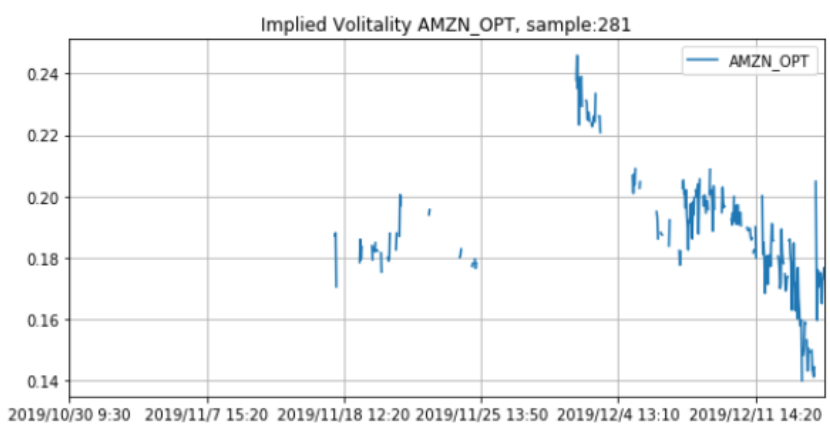

Figure 5 AMZN option implied volatility

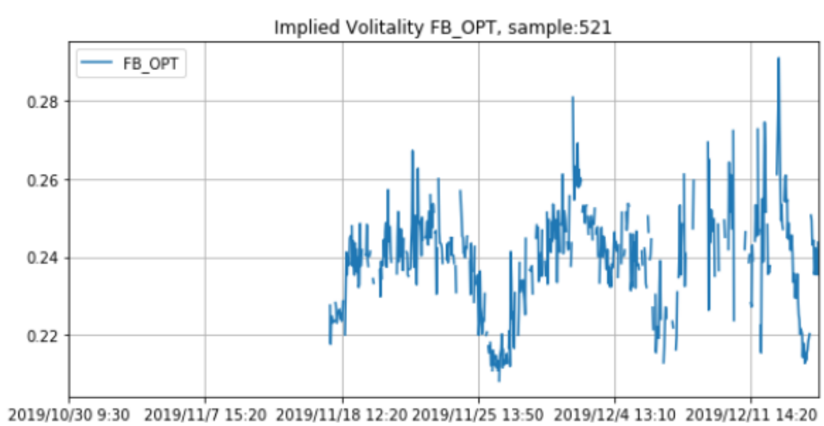

Figure 6 FB option implied volatility

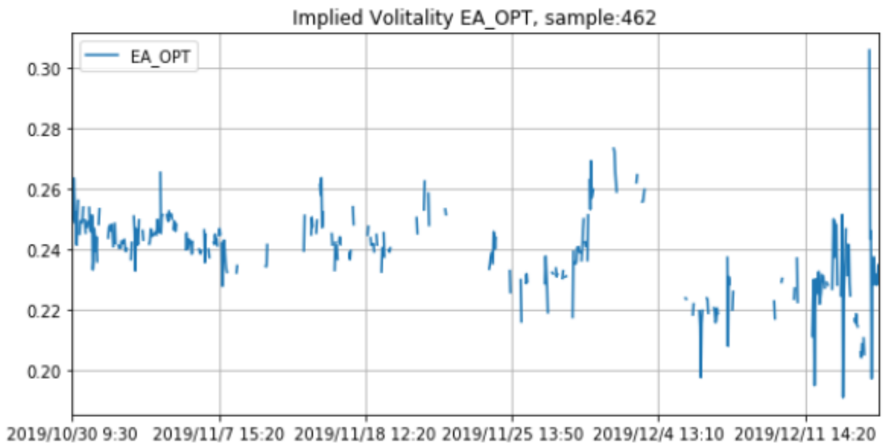

Figure 7 EA option implied volatility

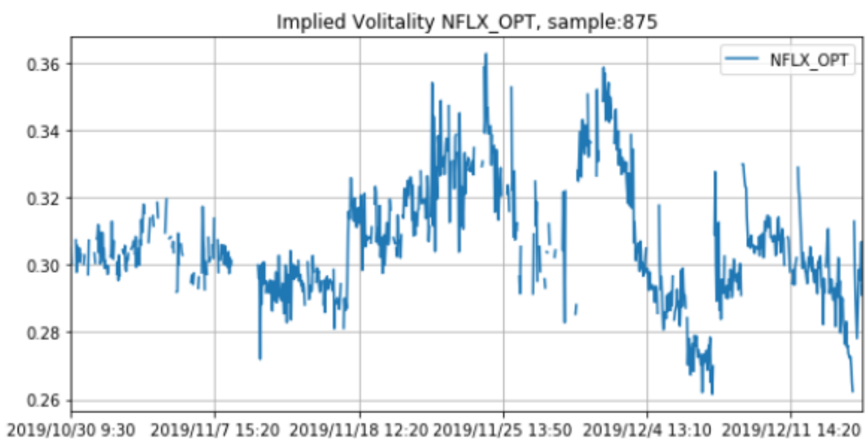

Figure 8 NFLX option implied volatility

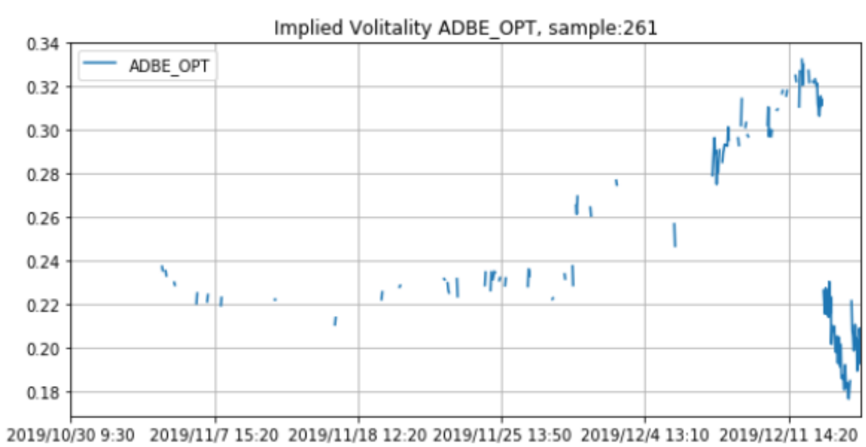

Figure 9 ADBE option implied volatility

The next step is to compare the forecasted volatility with the implied volatility. Although we tried 4 models for each stock in the previous section (ARCH, GARCH, IGARCH and TGARCH), we selected the best model for each option. We choose IGARCH for EA and FB, and ARCH for NFLX. The plots of implied volatility and forecasted volatility are shown below.

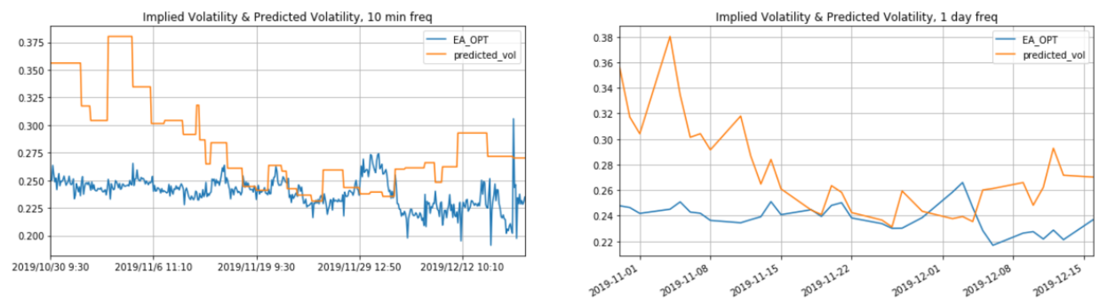

Figure 10 EA implied volatility with estimated volatility

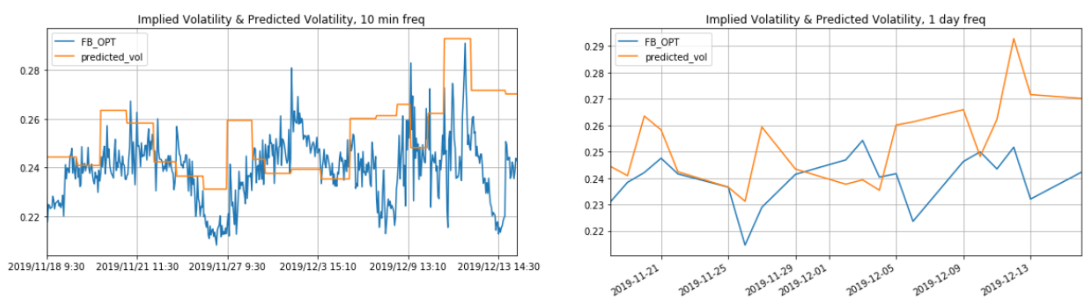

Figure 11 FB implied volatility with estimated volatility

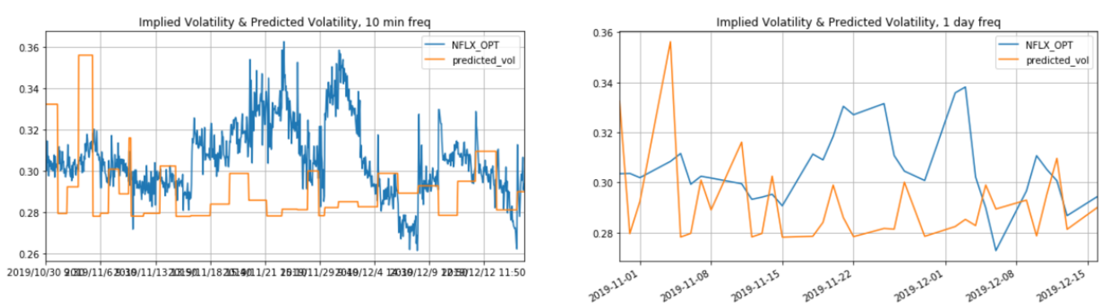

Figure 12 NFLX implied volatility with estimated volatility

In the left graph, we plot the daily forecasted volatility and the 10 minutes implied volatility. In the right graph, we plot the daily forecasted volatility and the daily average implied volatility. We can conclude from our graphs that our model make sense since the forecasted volatility differ not too much from the implied volatility.

5. Backtesting & Strategy Assessment

Our strategy is based on the Black-Scholes formula. According to the formula, we compare the implied volatility with the volatility predicted by our model. If the model indicates that the market overvalued the volatility, then we sell the call option. If the model indicates that the market undervalued the volatility, then we buy the call option. Every time we enter a trade, we hedge our portfolio by trading underlying stocks to make the portfolio delta neutral, thus we only make money by trading volatility.

Consider time T, if we do not hold any position at T, we just observe the implied volatility and our model's volatility. If the implied volatility is lower, we buy 1 option and sell delta shares of underlying stocks. We hold this portfolio until time T2 that the implied volatility first bigger than our model's volatility. In contrast, if the implied volatility is higher, we sell 1 option and buy delta shares of underlying stocks. We hold this portfolio until time T2 that the implied volatility first smaller than our model's volatility. The implied volatility we use is actually the average daily implied volatility since we need to downsample our frequency of data to daily basis.

We trade on EA, FB, NFLX and their call options. The test period is between 10/30/2019 and 12/12/2019. There are altogether 14 trades, with an averaging return 0.333. So, we conclude that our strategy works.

```
cumulative return of the strategy: 4.663
average return of the strategy: 0.333
total trade: 14
```

Figure 13 The result of backtesting.

6. Conclusion

Zero autocorrelation, mean reverting trend and extra kurtosis of the return could be effectively described by our models, while the autocorrelation of square return required by models is not significant in data. Max Likelihood Estimation along with nelder-mead optimization method and Schwartz Bayesian Criteria works fine on data set, providing us with useful information for trading.

We compare the volatility predicted by our model with the implied volatility of the call options and design a trading strategy which uses the difference between 2 volatilities. Although the strategy is a naive one, we got positive returns on it. This result proves that the volatilities suggest by our model is sometimes more efficient than the market and we can take this advantage to make a quite stable profit. One further improvement we can do is to do dynamic hedging when we open a position. Also, we can consider hedging gamma. These improvements are both indicated to enable a pure gain on volatility.

Appendix: Instruction on Jupiter Notebooks

(All the coding is in the *coding* file)

Data Exploration.ipynb: explore the characteristics of data.

Model_EA.ipynb: build time series models for EA.

Model_FB.ipynb: build time series models for FB.

Model_NFLX.ipynb: build time series models for NFLX.

Strategy _ Backtest.ipynb: backtest strategy based on the selected models and estimated volatility.