Naïve Bayes Classifier And Profitability of Options Gamma Trading

HYUNG SUP LIM
hlim2@stanford.edu

Abstract

At any given time during the lifetime of financial options, one can set up a variety of delta-neutral positions. In essence, delta-neutral volatility strategies attempt to make profit from correct prediction about the two components of options price: the mean reverting properties of implied volatility and the change of intrinsic value due to underlying price change. In this project, we focus on the latter, the intrinsic value, so we will assume constant volatility across the project. Delta neutral portfolio's sensitivity with respect to underlying is measured by the Greek letter delta and gamma, the first and second derivative of options value with respect to underlying price, respectively.

Volatility traders who take long gamma position bet that market would move fast enough to cover the cost of holding their delta neutral position. If one expect that market is going to stay constant for a certain amount time, then one might want to take short gamma position and collect the decreasing time value of options as profit. Thus, in delta-neutral volatility trading, it is crucial to predict whether or not the underlying market is about to move significantly anytime soon. We will test the profitability of short gamma strategy, so we want to answer the question: is it safe to short on options to collect time value now, or is the market swing imminent so short position on options will result in incurring loss? In this project, we apply Naïve Bayes classifier to KOSPI 200 index options data to classify each day's underlying price data into success or failure based on profitability of short gamma strategy.

I. Introduction

Motivation

The stock price movement is generally believed to be unpredictable. True news is random, and the market is generally efficient in reflecting new information into its price. Assuming the semi-efficient market, there might exist intelligent investors who have the information which has not yet been fully reflected in the price. There might also be some investors who are equipped with fundamental analysis and profound insights into the market. They will all try hard to make a directional bet on price. They would take a long position as they hope the price goes up or a short position vice versa. However, the introduction of financial derivatives, especially options, transform the rule of the game. Under the cloud of the directional uncertainty, financial derivatives allows the freedom from long or short only strategies and opens to door to variety of strategies that are irrelevant of directional moves. Then, where does the uncertainty part come in when it comes to delta-neutral options portfolio?

New Rule of The Game: Market Direction Does Not Matter

The uncertainty now comes from two sources: the implied volatility and gamma. The implied volatility is a portion of options contract price that we can consider as measure of investors' fear of uncertainty within the market. Gamma is the 2nd derivative of options price with respect to the underlying price. Equivalently, gamma is the first derivative of delta, and delta is the first derivative of options price with respect to underlying. In volatility strategies, we usually set up our options position so as to minimize our delta. So, our delta neutral position would be insensitive to small changes of underlying price.

However, our gamma is not zero, so large movement of underlying will result in change the value of our portfolio. Delta-neutral but positive gamma portfolio would generate profit if the underlying price moves significantly. However, positive gamma is not free. Positive gamma inevitably means negative theta. Theta is the 1st derivative of the options price with respect to the time. Thus, a positive gamma and negative theta position means that we might gain if the underlying price moves to either direction significantly, but we lose money each day if the underlying stays within some range. After all, the bet is not about whether or not the price would go up or down but about whether the underlying price would move significant enough to pay off our theta cost to hold that portfolio.

Effect of Dimension Reduction: Target Variable Is Now Just Binary

Suppose that we had to make a detailed prediction about price in the future, so response variable is multivariate. For example, if we had to predict not only the price would go up or down, but also how much it will go or down and how long it will take to reach that prediction price. Then, it would be very challenging, and we can expect the test error rate of our model to be very large. What if we do not have to predict the duration such that the dimension of time does not matter? The challenge of prediction would be easier. Now, suppose further that we do not have to predict how much it would move, and only be asked to predict whether it would go up or down. Then, it is now easier again. Similarly, in the option volatility trading, we only need to predict whether or not the price would move enough or not. Direction does not really matter. Hence, the introduction of volatility strategy has the effect of reducing the dimension of response variable such that it makes our prediction less challenging, if not easy.

II. Data

We use KOSPI 200 Options data. KOSPI 200 Options are based on KOSPI200 index, which is a market capitalization weighted index that consists of 200 blue chip stocks listed on the KRX stock market. KOSPI200 options is the most actively traded index options in the world today, so the options price is very efficient. The time series data ranges from Sep 1 2009 to Dec 5 2014, approximately 5 years of price data.

Top 20 Equity Index Futures & Options Contracts
Rank Contract Index Multiplier Jan-Dec 2011 Jan-Dec 2012 % Change

*1 Kospi 200 Options, KRX * 500,000 Korean won 3,671,662,258 1,575,394,249 -57.1%*
2 S&P CNX Nifty Options, NSE India 50 Indian rupees 868,684,582 803,086,926 -7.6%
*3 SPDR S&P 500 ETF Options ** N/A 729,478,419 585,945,819 -19.7%*
4 E-mini S&P 500 Futures, CME 50 U.S. dollars 620,368,790 474,278,939 -23.5%
5 RTS Futures, Moscow Exchange 2 U.S. dollars 377,845,640 321,031,540 -15.0%
6 Euro Stoxx 50 Futures, Eurex 10 euros 408,860,002 315,179,597 -22.9%
7 Euro Stoxx 50 Options, Eurex 10 euros 369,241,952 280,610,954 -24.0%
8 S&P 500 Options, CBOE 100 U.S. dollars 197,509,449 174,457,138 -11.7%
9 Sensex Options, BSE 15 Indian rupees 383,543 148,314,519 38569.6%
10 Nikkei 225 Mini Futures, OSE 100 yen 117,905,210 130,443,680 10.6%

$$M^{50\times 60}, where\ M^{i\times j} = ith\ value\ of\ KOSPI\ 200\ \text{Index}\ within\ jth\ expiration$$

During that period we had 60 options expiration dates. For each 60 options expiration, the price data is partitioned such that each partition has 50-days time series data We set up a 50x60 matrix, in which i-th row represents the ith-date within options life time, and j-th column corresponds to j-th expiration options.

III. Data Preprocessing and Target Variable

We want y to indicate whether or not it is worthwhile to take a delta neutral position at time t. Because we are focusing on gamma aspect of volatility strategies, we will assume that volatility is constant throughout the project. Hence, we will define y as the following.

$$\begin{cases} y = 1\ if\ \ \delta\Delta P\ + \left(\frac{1}{2}\right)\gamma\Delta P^2 < \theta : if\ short\ gamma\ position\ makes\ money \\ y = 0\ if\ \delta\Delta P\ + \left(\frac{1}{2}\right)\gamma\Delta P^2 > \theta : if\ short\ gamma\ position\ loses\ money \\ , where\ \Delta P\text{: } change\ of\ underlying\ price, \\ \gamma\text{: } gamma\ sensitivity\ (2nd\ derivative\ of\ options\ price\ with\ respect\ to\ underlying) \\ \theta\text{: } is\ the\ theta\ sentitivity(derivative\ of\ options\ price\ with\ respect\ to\ time) \end{cases}$$

The definition above uses quadratic approximation. We chose y to be 1 in order to indicate that short volatility strategy is successful. If y is 1, then it means we collect the time decay of options as our profit, which is net positive. Since our position is delta neutral, our portfolio's value does not vary much due to $\delta \Delta P \approx 0$. Hence, most of portfolio value variation comes from $(\frac{1}{2})\gamma \Delta P^2$ assuming constant volatility. We want this to be small and time decay to be large to maximize our profit.

We pre-processed our KOSPI 200 data to match above definition of target variable. Before the pre-processing, y variable is just raw KOSPI 200 Index value. Data preprocessing for target variable followed the following pseudo code. And we obtained the response variable of randomly repeating 1's and 0's from double type price data.

```
pseudo code
for each time series entry
  find the delta minimizing strike prices for both ITM and OTM at time t
  find delta gamma theta of portfolio with strikes found above at time t
  compute the threshold value with quadratic approximation and time decay
  if time value is greater than the threshold ,
    then store y_{t+1} = 1, y_{t+1} = 0 otherwise.
```

IV. Model: Naïve Bayes Classifier

$$\text{Naive Bayes:} \quad P(Y = y_k \mid X_1 \dots X_n) = \frac{P(Y = y_k) \prod_i P(X \mid Y = y_k)}{\sum_j P(Y = y_k) \prod_i P(X \mid Y = y_j)}$$

In this project, we focus on the gamma component of the delta-neutral options portfolio. On the other hand, the implied volatility of options portfolio is widely believed to be clustered as referred to as volatility clustering. It seems intuitive that large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes. Applying similar concept to gamma aspect of options, we propose a hypothesis that fast moving period is clustered together, and slow moving period is also clustered. Based on this idea, we include some feature variables which might have some predictive power. We will include various features into our model, and they are discussed that later.

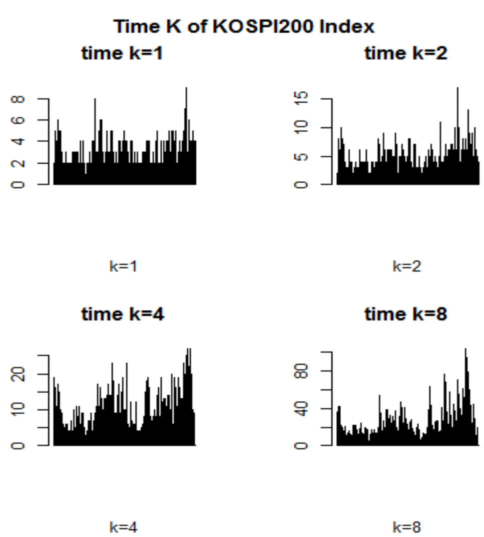
For our model, we used Naïve Bayes Classifier with pre-processed y variables as discussed in previous section and features are chosen in a way that are related to the clustering characteristic of underlying prices.

V. Features

1. Introduction of New Variable: Price Rapidity

Essentially, we are interested in predicting how far the future price will be away from the current price, regardless of direction. Consider the quantity:

|P_(future)-P_current |



This quantity is independent of direction, and we want this to be small in our short gamma strategy. Because of the trade-off relationship between this quantity and the time decay, we are naturally interested in how long it takes for the price to change by k units. In other words, how many days do we have to wait to see k units of price change? In this idea, we defined a variable time-k, which is the number of days it took since the time when the price was different k units. It goes retrospectively such that it traces back to the past from now to find the value time-k. For example, consider the following price change. 201 → 195 → 200 → 201 → 202 → 200(now). The current price 200. If we want to find the value time-3, it will be 4 because the first day from now for the price to be different by 3 is t4. |195-200|>3. We computed the time-k with simple algorithm, and we obtained above figure.

As you can see, the sporadic spikes at the figure of time-k strengthens our hypothesis that the volatile and sluggish periods are clustered together. Note that large value of time-k implies that it has taken long time to differ by k. So, large time-k means less volatile market, vice versa. Thus, the inverse of this quantity would be rapidity of market. We defind the rapidity of our market:

$$Rapidity\ _k = \frac{1}{Time_k}$$

We included this quantity in our Naïve Bayes classifier as our feature variable assuming the backward rapidity of the market might have some explanatory effect on the gamma and theta trade-off dynamics.

2. Other Features

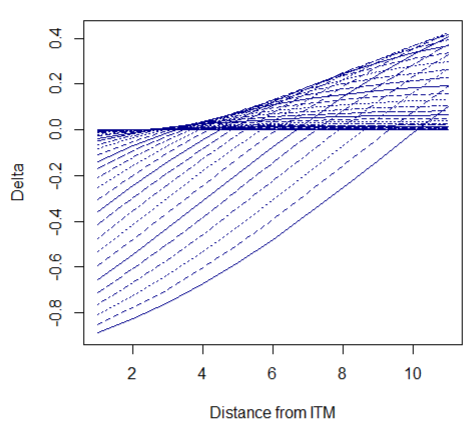
Again, we are choosing our features that we think might help explain the clustering characteristic of our response variable. So, we included lagged y values up to order 10. Previous data itself might have some forecasting power on response. We hope that some variance of response variables can be explained by their own past values. We also included the number of days that showed the same type of response variable as the current one. For example, if a series of response showed the patter of 0-> 1->1->1(now), then we included 3 as our feature to indicate that the series has shown 1(success) three days in a row. We think that if a previous response variable showed success, for an example, then it is more likely to show success than failure. However, if the series shows many consecutive success in the past, it is very likely to reverse the pattern, and it alternates this pattern. We did the same procedure for 0(failure), too. We also added days since last success and last failure, regardless of consecutive pattern.

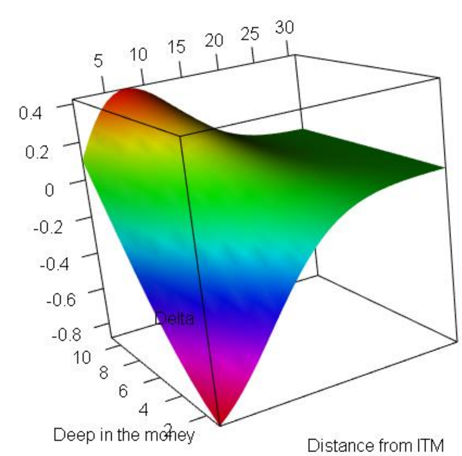
Summary of Features	
$lag_1 \sim lag_2$	*Lagging response variable with order 1~10*
$Csec_1, Csec_0$	*The number of consecutive 1 and 0* *If the current y value is different from previous, then this value is zero.*
$toLast_1, toLast_0$	*The number of days since last 1 and 0, regardless of current value of y*
$Rapidity_1 \sim Rapidity_{10}$	*Rapidity of price as defined above*

VI. Delta Neutral Volatility Trading Strategy: Delta-Neutral Portfolio

Delta Neutral Position: In order to implement an options volatility strategy, we need to set up a delta neutral portfolio. A variety of delta-neutral portfolio is possible, but here we set up the simplest kind of delta-neutral position for our trading algorithm's execution efficiency.

Strategy	
$\begin{cases} \text{buy or sell one ITM and sell or buy two OTM} \\ V = (\pm)1 \times ITM + (\mp)2 \times OTM \\ \text{choose ITM and OTM such that } argmin_{itm,otm} \dfrac{dV}{dp} \\ \text{, where ITM is any lower strike call option} \\ \text{OTM is any call with higher strike than ITM's strike} \\ \dfrac{dV}{dp} \text{ is portfolio's sensitivity to underlying price} \end{cases}$	For each time, we want to minimize $\lvert D_{pt}(t) \rvert$ (delta of a portfolio at time t) $, where\ D_{pt}(t) = N\left(\frac{1}{\sigma\sqrt{T-\tau}}\left[ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-\tau)\right]\right)$ N is Normal Distribution's CDF S = underlying price, K = strike price $T - t$ = time to maturity, σ = volatility, r = risk − free interest rate consistently rebalance the portfolio to minize delta



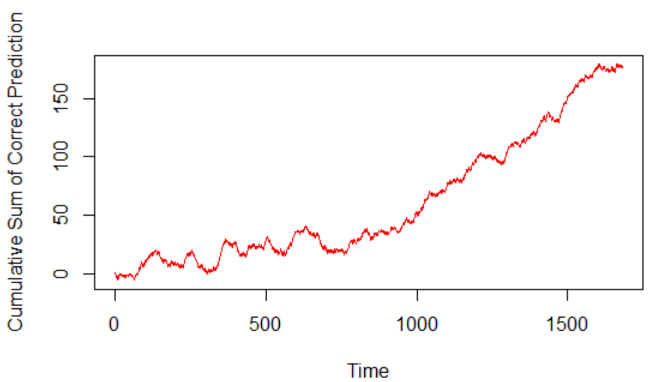


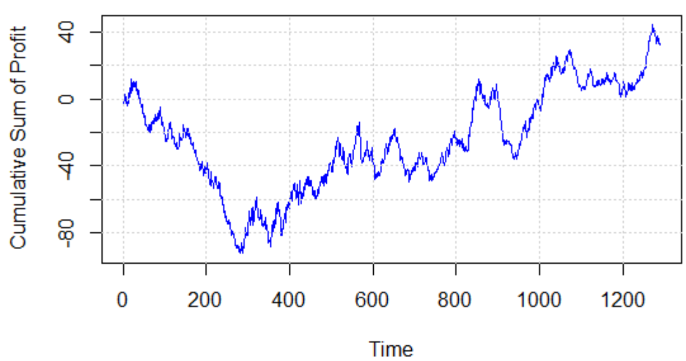
VII. Results and Trading Performance

In our model, the classifier learned in a cumulative fashion. Notice that price data is not a random permutation of variables. Every data entry is ordered, and the specific order might have some features related to response. Hence, not all of the data is available from the beginning. In the beginning, we only have a small length data, but as we go through, we accumulate the historical data such that we have larger data set. For example, among the 60 expirations of time series data, the 1st expiration data was used for training purpose only. Using the trained model from 1st expiration, we made prediction about 1st response in the 2nd expiration. For the next response in the 2nd expiration, we accumulated the whole 1st expiration data and the first response in 2nd expiration. We trained our classifier again and used it for the prediction about the next response. Hence, as we go across the whole time series, our classifier will have learned from the larger data set. Our simulation results supports our hypothesis that the classifier would perform better with larger training data set. For the result vector, we stored +1 if the classifier predicts correctly, and -1 if otherwise. And we obtained the cumulative summation of the vector, and the plot is shown below.

As we can see the figure below, the learning rate of our classifier does not seem very fast in the beginning period. However, it starts learning at a faster rate after about 1000th data. This result is very promising because if we collect more and more data such that our time series spans a longer time period, our classifier would perform even better.

Trading simulations is performed in the following way. If our classifier predicts that y variable is 1 in the next day, we set up short gamma delta neutral position assuming constant volatility. And we unwind the position next day, and we keep repeating this process. The trading simulation result seemed very impressive to us because it is possible to generate positive profit in the long run even though there seems to be some cost in the beginning phrase where the classifier need some trial-and-error experience.





VIII. Future

We only used some transformation of response variable such as lags and the number of days since last success. If we could include some other variable that are related to the market such as interest rate or some other regularly announced economic indicators, we might also be able to improve our model. In future, we want to use other linear classifiers such as perceptron and SVM with same formulation of training data. We also want to trade with this model in the real options market with real money so that we can learn what other factors affect the trading performance.

References

Mandelbrot, B. B., The Variation of Certain Speculative Prices, The Journal of Business 36, No. 4, (1963), 394-419
John Fox & Sanford Weisberg., An Appendix to An R Companion to Applied Regression, Second Edition (2011)
Will Acworth., FIA Annual Volume Survey: Trading Falls 15.3% in 2012, 24 (2012)
Black, Fischer; Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". Journal of Political Economy 81 (3): 637–654. doi:10.1086/260062. [1] (Black and Scholes' original paper.)
Merton, Robert C. (1973). "Theory of Rational Option Pricing". Bell Journal of Economics and Management Science (The RAND Corporation) 4 (1): 141–183. doi:10.2307/3003143. JSTOR 3003143. [2]
Hull, John C. (1997). Options, Futures, and Other Derivatives. Prentice Hall. ISBN 0-13-601589-1.