High Frequency Strategy Using Statistical Arbitrage

Group Members

Shu Wang, Dan Zhu, Kun Qian

Abstract

High Frequency Trading is a vital part of stock trading at present. There exist various kinds of trading strategies designed for multiple purpose. One named statistical arbitrage is the strategy profiting from the volatility of the stock market while creating a portfolio immunized from the risk of price changes. We studied relevant papers and produced the realization trading program of this method, implemented it on the US stock market and made some modification of the parameters. The best result shows that this high frequency strategy obtains small returns with low volatility.

(High Frequency Trading; Statistical Arbitrage; Intraday Volatility)

1. Introduction

High frequency trading (HFT), as a famous trading method in financial markets nowadays, actually come up since the end of last century in 1983, shortly after NASDAQ introduced a purely electronic form of trading. Of course the term itself is not limited to modern financial market, and thus can be traced back to earlier, but that's out of our consideration. Until now there is still not a universal definition of HFT, but Narang gave us a quite reasonable one:

"High-frequency traders (a) require a high-speed trading infrastructure, (b) have investment time horizons less than one day, and (c) generally try to end the day with no positions whatsoever."

High frequency trading is an innovation in financial markets. With the speed, much faster than human being to execute, high frequency trading can receive information and complete orders in milliseconds by using powerful computers and advanced programs. On the one hand, high frequency trading is beneficial for stock investors since it can lower trading costs and raising trading volume and liquidity. On the other hand, it poses potential systemic risk and increases stock volatility. Moreover, high frequency trading can help traders to better and faster process relative information.

There are several kinds of HFT such as passive market making, arbitrage, structural and directional, etc. Thus those companies which are involved in these trades is the main market participants of HFT. Hollis (2013) wrote that "the major U.S. high frequency trading firms include Chicago Trading, Virtu Financial, Timber Hill, KCG, IMC, Tradebot and Citadel LLC." With more and more high frequency trading strategies been used in the market, the profit went down dramatically. In the words of Cookson (2013), in the United States, profits from high frequency trading has been dropping from an estimated peak of 5 billion in 2009, to about 1.25 billion in 2012. HFT is also a common thing in futures market. But the total volume of the HFT is a rather vague number.

In the early 2000s, high frequency trading only hold fewer than 10% of equity orders. However, its proportion increase rapidly afterwards. Duhigg (2009) stated that "trading volume grew by about 164% between 2005 and 2009 for which high frequency trading might be accounted from data from the NYSE." High frequency trading contributed 60% to 70% of U.S. trading during 2009 to 2010, but the percentage has decreased in the last few years. According to the Aite Group, HFTs in 2013 account for a little more than half of global equity volumes, about the same percentage of futures volumes, and about 40 percent of currency volumes. In equities specifically, Aite estimates that HFT's share of trading is highest in the United States (again, a little over half), more than 40 percent in Europe, and almost 20 percent in Asia. With the help of high frequency trading growth, other market investors can trade cheaper since bid-offer spreads are narrowed.

2. The Intraday Statistical Arbitrage

In this section, we first review theories for a quite straight forward trading strategy, which uses two similar equal-weighted portfolios of most liquid stocks on the market only with different rebalancing frequency. The main idea of our algorithm is referenced from *Robert and Cary (2007)*. Then we will explain how we apply this strategy in detail using quantopian.com.

Rationale of the Strategy

Robert and Cary (2007) states that high-speed trading strategies, which is referred to as statistical arbitrage, is a powerful tool to attain high information ratio.

High-speed trading strategies similar to market making have putatively been used by hedge funds in recent years. Market makers in financial markets generate profits by buying low and selling high over short time intervals. These portfolios act as market makers by selling on upticks and buying on downticks. The idea of sell high buy low is one of the most basic trading strategies, but the circumstance in which one can exactly define high and low is that the market has mean-reverting property. This is exactly one assumption the strategy based on. Also to avoid risk from difference between close price and next open price, the strategy is an intraday strategy, i.e. the balance will be cleared every day before close, everyday gains and losses can be recognized then.

For a long-only fixed weights strategy, the high-speed trading portfolio, which rebalances more frequently, attains an excess growth rate which is proportional to the relative variances of the stocks and is non-negative. This can be verified via following definition of logarithmic return of an n-stock equal-weight portfolio(P):

$$\mathrm{dlogP} = \sum_{i=1}^{n} p_i dlogX_i + \gamma_p^* dt$$

where $\gamma_{\mathrm{p}}^* = \frac{1}{2}\sum_{i=1}^{n} p_i \sigma_{ip}^2$, σ_{ip}^2 is the variance of $\mathrm{X_i}$ relative to the portfolio

It can be concluded the second term gives us an extra return while the first term can be cancelled out because we are longing and shorting exactly the same portfolio at the beginning of every trading day. Thus, the hedging strategy, with high-speed portfolio and buy-and-hold portfolio at opposite sides, can attain a riskless profit (relative to the benchmark).

The empirical analysis in the paper shows that for the universe of large capitalization U.S. stocks, even quite naive techniques can achieve remarkably high information ratios. The authors plot the daily variance versus the rebalance frequency, finding that estimate of daily variance decreases as the frequency decreases from a 1.5-minute sampling interval to a 390-minute interval. Thus the strategy is to long the 1.5-minute rebalancing portfolio, while short the 390-minute one. By the definition we listed at the beginning of the report, this is exactly a HFT strategy. Below we'll explain how the computer program works.

There is one thing worth notice, when the market has obvious trend, the strategy then might face huge losses. Before you rebalance, i.e. at the beginning of any trading day, same weight on stocks makes your position protected. But once you rebalance, it happens that the sale of stocks on upticks can't afford to buy those on downticks, once the value difference between the two portfolios exists, then some position must be at risk even you rebalance the portfolios back to equal weight.

Another minor problem is the conflict between the nature of the statistical arbitrage and the profitability of it. The strategy requires a rebalance of its portfolio in 1.5 minutes to keep it a constant, equal-weighted portfolio, but the article itself has mentioned that in reality constant portfolio is hard to achieved. Actually $100 can't buy even a few shares, you need probably millions to go through the technical problems such as rounding of the number of shares and reach a relatively constant, equal-weighted one. But large quantity of trading provides you with market impact i.e. liquidity issue, then, which no wonder will decrease the profit. Besides the transaction fee and other minor problems are ignored by the strategy too.

3. Implementation on US stock market

We implemented our high frequency trading strategy on US stock market during year 2016 and 2017. The strategy is to long the *f*-minute rebalancing equal-weight portfolio and short the 390-minute rebalancing portfolio (intraday buy and hold portfolio). In practice, we calculate the positions differences between the high-frequency portfolio and the low-frequency portfolio, and do the trading to achieve the positions.
In the empirical analysis, we take several experiments, modifying some parameters and backtesting to find the optimal combination. The parameters are: 1) method of stock choosing 2) equal-weight versus market capitalization weighted; 3) frequency of rebalancing. Meanwhile, we set up the stop-loss line and enlarge the positions 10 times to control the risks and increase the return.
We backtest our strategy on Quantopian (https://www.quantopian.com). All experiments are based on a date ranging from 01/01/2016 to 02/28/2017, slippage and commission are set as default by Quantopian to resemble the real market.

1) Method of Stock Choosing

One observation is that when the return of the benchmark resembles significant trend and low volatility, the return of our portfolio is lower, while a higher return is expected when more volatile is included in the benchmark return. This phenomenon can be explained by the structure of our portfolio. We sale securities whose price goes up and buy those whose price goes down. However, when the market continually goes up (or done), our portfolio will sale at the beginning at a relative lower price and buy at the end of the day at a higher price, which will draw down the return of our portfolio. The same result happens when market goes down continually for a period.

Thus, our hedging portfolio benefits most in a high volatile market while performs badly when market observes a significant trend.

To benefit from this observation, we do three filtrations through pipeline to our pool of stocks and select those with high volatility after a selection of high dollar volume and finally pick the top stocks with high market capitalization. After several changes to the sample size of each filter, we end up with a reasonable and the best performed one: rank dollar volume from high to low and pick the top 500 stocks, then rank volatility from high to low and pick the top 100 stocks, at last rank market capitalization from high to low and pick the top 50. High volume ensures there are enough liquidity in the pool of stocks, so does high market capitalization.

We list a few back-test results in the table below, keeping other parameters unchanged:

Table 1. Effect of volatility selection on return. Frequency: 2/390, equal-weighted. The time window is 06/01/2016-11/30/2017.

Test number	parameter	Return (%)
2	dollar volume(200), volatility selection (100), market_cap (50)	0.07
3	dollar volume(500), volatility selection(100), market_cap(50)	0.12

For all three tests, benchmark return is 18.46%, beta is 0, portfolio volatility equals 0. We increase the sample of dollar volume in test 3 to contain more fluctuation in our pool of stocks, thus the return is even higher. Parameters for test 3 are chosen in latter experiments.

2) Equal-weight versus Market Capitalization Weighted

Let $X_i(t)$ be the market capitalization of stock i, the market weight of stock i at time t is

$$\mu_i(t) = \frac{X_i(t)}{X_1(t) + \cdots + X_n(t)}$$

where $n = 50$ is the number of stocks in our pool. For a parameter $0 \leq p \leq 1$, the diversity-weighted portfolio has weights

$$\omega_i(t) = \frac{\mu_i^p(t)}{\sum_{j=1}^{n} \mu_j^p(t)}, \qquad i = 1, \ldots, n$$

To change the portfolio from equal weighted to market_cap weighted, simply change from $p = 0$ to $p = 1$ in the above formula.

Table 2. Effect of weight on return. Frequency: 10/390.

Test number	Weight	Return (%)
4	Equal weight	-0.06
5	Market_cap weight	-0.12

The result turns out to be worse for market_cap weighted portfolio. This is because market_cap weighted portfolio put more weight on the large_cap stocks, while these stocks normally resemble low volatility, with more weight on them, the combined portfolio will have less volatility compared to equal-weighted portfolio. Therefore, by what we have discussed in part 1), the final return will be lower. In the following part of our report, we will continually use equal-weighted portfolio.

3) Frequency of Rebalancing

As our strategy is hedged between a high frequency portfolio and a buy and hold portfolio, thus in theory, the more frequent the former strategy is, the more volatile it can capture and more returns is generated. We test this idea by changing the frequency of rebalance from 10/390 to 3/390 to

1/390, in other words, rebalance the high frequency portfolio from every 10 minutes to every 3 minutes to every 1 minutes, and keep the buy and hold portfolio's frequency unchanged, that is rebalance only when market open and market close. The result is shown below:

Table 3. Effect of frequency on return.

Test number	Frequency	Return (%)
6	10/390	-0.07
7	3/390	-0.04
8	1/390	-0.08

Test 7 showed what we expect, a lower return compared to test 6 as frequency increases. While test 8 had an unexpected worst return. This is probably due to order cancellation of unfulfilled orders as we rebalance too frequently, some orders did not get enough time to be filled before the next rebalance time. Thus, open positions and actual transactions are lower as what we had desired, causing the final return to be lower than expected.

To make up for the loss when frequency equals 1/390, we can slightly increase the frequency to 2/390, the result turns out to be better. We will show the result in the last part of our report, which combined all changing parameters together and produced an optimal result.

4) Stop-loss Line

Figure 1 shows the returns of each day during 01/01/2016 – 02/28/2017. On March 15th, 2016, the return is -$17,702 (0.17%).

As is mentioned above, this strategy performs bad and may lose money when the price has a trend. This risk can be lowered by setting stop loss line, and once the return touches the line, we close all positions and stop trading in this day.

Table 4. Effect of stop-loss line in return.

Test number	Stop-loss return (%)	Return (%)
9	$-\infty$	-0.12
10	-0.03	-0.02
11	-0.01	-0.11

Comparing between Test 10 and Test 9, the stop-loss line increases the total return. Comparing between Test 11 and Test 10, setting a higher stop-loss line decreases the return. The reason is, the trading is terminated too early and then lose the opportunity of returns getting back. Finally, we pick the threshold as -0.03%.

5) Enlarge the Portfolio

We notice that the daily transaction is significantly low, which is about 1 million dollars, comparing to the total capitals (10 million dollars). That means, we did not fully utilize the capital and nearly 90% of the capital is just put in the bank and never be used during the whole trading. Therefore, we can enlarge our portfolio 10 times by enlarging positions of all assets 10 times and obtain a higher return.

Our best result

The bests result is obtained by implementing our strategy during 06/01/2016 – 11/30/2017, with other parameters shown below. Since high frequency trading strategy is best used by mutual funds which is commission free, we set the commission fee as 0 while backtesting.

Table 5. Parameters choosing

Stock choosing	dollar volume(500), volatility selection(100), market_cap(50)
Weight	Equal-weighted
Frequency	2/390
Stop-loss line	-0.3%
Enlarge ratio	10

Table 6. Results.

Total returns	1.38%
Alpha	0.01
Beta	0
Sharpe ratio	1.06
Volatility	0.01
Max drawdown	-0.66%

4. Conclusions

High frequency trading obtains small returns with low volatility under the commission free situation. Since we hedged the high frequency portfolio by the low frequency portfolio, the strategy is beta-neutral. Meanwhile, the positions on assets are actually not large, resulting to the low daily returns (approximately 0.01%) and hence the total return is not large. Since the Sharpe ratio is positive and the volatility is extremely small, we can still obtain (approximately) riskless profit from the statistic arbitrage and the intraday volatility.

Further research can be established from several aspects: 1) use some other portfolio weights instead of equal weights or market cap weights; 2) implements some more effective methods to filter stocks to obtain more returns from volatility as well as lower the risk of crash.

References

Duhigg, Charles (July 23, 2009). "Stock Traders Find Speed Pays, in Milliseconds". New York Times. Retrieved Sep 10, 2010.

Hollis, James E. (Sep 2013). "HFT: Boon? Or Impending Disaster?". Cutter Associates. Retrieved June 29, 2015.

Cookson, Clive (May 12, 2013). "Time is money when it comes to microwaves". Financial Times. Retrieved May 12, 2013.

Securities and Exchange Commission, & Securities and Exchange Commission. (2010). Findings regarding the market events of May 6, 2010. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. Chicago.

Aldridge, Irene (2009). High-frequency trading: a practical guide to algorithmic strategies and trading systems.

Vol. 459. John Wiley and Sons.

Gomber, Peter et al. (2011). “High-frequency trading”. In:

Narang, Rishi K (2013). Inside the Black Box: A Simple Guide to Quantitative and High Frequency Trading.

Vol. 883. John Wiley & Sons.

Phillips, Matt (2010). “Nasdaq: Here’s our timeline of the flash crash”. In: The Wall Street Journal.