Hurst Quant Investment Strategy Implementation with Python

Alan Moreira

Yihua NI

Guanyu ZHANG

Ruoyi Chen

SUMMARY

The Hurst exponent, originally invented to predict the flooding in the filed of hydrology, is an excessive used statistical technique to detect and evaluate the amount of persistence or randomness in financial markets times series data. Long-range dependence and its presence in the financial time series has been wildly discussed and it is believed that the historical trend of the stock are informative for its future performance. Therefore, our strategy used the Hurst exponent as the signal since it can capture the self-similarity of the stocks.

In this project, we applied the R/S method to evaluate the value of the Hurst exponent as it is well developed and simple and direct to understand. Through analyzing the excessive return of 50 random stocks, we built 10 investment portfolios based on the value of Hurst and the sign of the Beta. The portfolios were constructed on a rolling basis and updated at the end of each year. Based on our study, the investors should long the portfolio with Hurst which is largest deviation from 0.5 and positive Beta, and short the portfolios with the opposite performance on Hurst and Beta. By implementing our strategy, the investors can outperform the market. Our strategy has zero alpha thus there is no abnormal return and the returns of the strategy are compensation for the risk.

Another wildly considered trading strategy in the financial market is momentum. Both strategies are focusing on the historical performance of the stocks. However, the momentum simply assume that the past trend would be repeated in the future. Our strategy, on the contrary, not only considers the historical trend, but also capture its presence of that trend. Therefore, the Hurst exponent is more informative and comprehensive than the momentum strategy.

We estimate the alpha and beta using both the CAPM and FF3 model. Under both model, the Tstatistic of alpha for each portfolio are not statistically significant. In CAPM model, the beta varies from 0.3 to 1.3. And the Sharpe ratio shows the trend that the better the signal, the higher the Sharpe ratio. In addition, the appraisal ratio reveals the same trend. In the FF3 model, the tstatics of betas are within the range of (-1.96, 1.96) thus are statistically insignificant.

Although our strategy has performed well based our analysis, however, it has own limitations. As the market changes over time and the outperformed stock varies, our strategy requires the active management of the portfolios and frequent transactions. The expenses and fees attached can affect the overall return of the investors. Besides the cost, this strategy has the inherent risk as there are many methods to estimate the Hurst which deliver different results. Another risk is that our strategy only focus on the past while ignores current situation of the company. Also, our trading signal fails to capture the impact of the value of Beta thus the accuracy may be impaired.

In summary, the Hurst exponent reveals the long-range dependence and its presence of the historical data and thus can be used as the trading strategy. Although this strategy has some limitations, it also has valuable properties when compared with the momentum strategy in the similar category. In this project, this strategy was verified to have outstanding performance and could be considered as useful strategy of the investors.

We have illustrated our strategy in video: https://youtu.be/lobo8vAqBos

1. Signal: The Hurst exponent

1.1 Introduction:

The Hurst exponent, proposed by H. E. Hurst for use in fractal analysis, has been applied to many research fields. It has recently become popular in the finance community especially in quantitative trading.

The Hurst exponent is used as a measure of long-term memory of time series. It relates to the autocorrelations of the time series, and the rate at which these decrease as the lag between pairs of values increases. Studies involving the Hurst exponent were originally developed in hydrology for the practical matter of determining optimum dam sizing for the Nile river's volatile rain and drought conditions that had been observed over a long period of time. The name " Hurst exponent", or "Hurst coefficient", derives from Harold Edwin Hurst, who was the lead researcher in these studies; the use of the standard notation H for the coefficient relates to his name also.

In fractal geometry, the generalized Hurst exponent has been denoted by H or Hq in honor of both Harold Edwin Hurst and Ludwig Otto Hölder by Benoît Mandelbrot. H is directly related to fractal dimension, D, and is a measure of a data series' "mild" or "wild" randomness.

1.2 Methodology

This strategy combines Hurst exponent and Beta (the return growth rate) to decide whether to buy or to short a stock.

Hurst exponent is used to describing the long-range dependency of a time series. A number of estimators have been proposed in the literature. The oldest and best-known one is the so-called rescaled range ***(R/S)*** analysis popularized by Mandelbrot and Wallis and based on previous hydrological findings of Hurst. The long-range dependence phenomenon exists in lots of fields. Researchers found that the stocks performed well in the past tend to perform well in the future, and stocks whose return decreased in the past tend to continue to decrease in the future. In 1994, Peters applied Hurst exponent in the capital market. (Peters 1994) From then on, more and more people began to use Hurst exponent to capture the trend of stocks' returns.

1) Calculate the mean: $m_t = \frac{1}{t}\sum_{i=1}^{t} X_i$

2) Calculate the deviation series: $Y_t = X_t - m_t$

3) Calculate the cumulative deviate series Z: $Z_t = \sum_{i=1}^{t} Y_i$

4) Compute the range R: $R_t = \max(Z_1, Z_2, \ldots, Z_t) - \min(Z_1, Z_2, \ldots, Z_t)$

5) Compute the standard deviation S: $S_t = \sqrt{\frac{1}{t}\sum_{i=1}^{t}(X_i - \bar{X})^2}$

6) Calculate the rescaled range ***R(n)/S(n)*** and average over all the partial time series of n.

$$(R/S)_t = R_t/S_t$$

7) Repeat step1-step6 to calculate *R/S*, but use different t value (t=2,3…n). i.e. Calculate the R/S of the following series.

8) Calculate Hurst exponent.

Use ***lg(R/S)$_t$*** as the dependent variable and ***lgt*** as the independent variable to do a regression, the slope H is Hurst exponent.

The Hurst exponent is estimated by fitting the power law (shown below) to the data.

$$\mathrm{E}\left[\frac{R(n)}{S(n)}\right] = An^H \text{ as } n \to \infty$$

This can be done by plotting ***log[R(n)/S(n)] as*** a function of log n and fitting a straight line; the slope of the line gives H (a more principled approach fits the power law in a maximumlikelihood fashion).

1.3 Interpretation

(1) H=0.5 indicates a random series. (2) 0<H<0.5 indicates an anti-persistent series. (3) 0.5<H<1 indicates a persistent series. An anti-persistent series has a characteristic of "mean-reverting", which means an up value is more likely followed by a down value, and vice versa. The strength of "mean-reverting" increases as H approaches 0.0. A persistent series is trend reinforcing, which means the direction (up or down compared to the last value) of the next value is more likely the same as current value. The strength of trend increases as H approaches 1.0. Most economic and financial time series are persistent with H>0.5.

2. Sample and data cleaning

2.1 Sample selecting

At first we selected 30-year's monthly returns of these 50 stocks, but considering that Financial Crises in 2008 will affect the law of returns, we choose the recent 5 years' data.

2.2 Data cleaning

Before analyze the data, we should do the preparation work first. We use the rolling window function of Beta and Hurst index to show the performance of 10-level portfolios, which is based on the last 12 months' data, and then we test them with the new return and Sharpe ratio of the next month.

The Hurst construction we have already explained in last part, and the Beta construction uses the ***Polyfit*** function to get the relation between time and returns in last 12 months. After constructing the rolling window, we save and reload our rolling window in the CSV to uniform the format.

2.3 Portfolios construction methodology

To construct the 10-level portfolios, we should make sure the Signal, and it will help us to divide data into different level portfolios. Hurst index indicates consistency of the history returns, and Beta means the relationship between returns and this period of time. So if Hurst > 0.5 and Beta > 0, it means that the past return increased and it will continue to increase; if Hurst < 0.5 and Beta < 0, it means that the past return decreased but it will not continue to decrease, it will increase instead. Both 2 situations are what we want. However, if Hurst > 0.5 and Beta < 0, it means that the past return decreased and it will continue to decrease; if Hurst < 0.5 and Beta > 0, it means that the past return increased but it will not continue to increase, it will decrease instead. Both 2 situations are what we don't want.

Under the theory above, we could then quantify the Signal. If Hurst>0.5 and Beta<0, or Hurst < 0.5 and Beta < 0, we could express the Signal as ***abs('Hurst' - 0.5)*** , which measures the level of consistency. The higher this level of consistency it gets, the higher value of the Signal we will get. However, if Hurst > 0.5 and Beta < 0, or Hurst < 0.5 and Beta > 0, we could express the Signal as ***-abs('Hurst' - 0.5).*** Because this situation won't bring us with great profit, so when the Hurst is far away from 0.5, the Signal will be smaller.

Next, we could get our 10-level portfolios by dividing by the value of Signal. To realize it, we use ***Qcut*** function to get 10 different level's signals, from the lowest to the highest, and then divide each 5 stocks into a group, which has similar Signal values.

Consequently, we could use simple weight times stock returns then divide the sum of the total weights to get the weighted returns. Then we could get the signals using the 12-months rolling window 's data, and get the 10-level portfolio return on these signals.

Finally, we could get the rolling window's results of returns each month to test our thoughts. The Signal9 indicates the most successful portfolio with 5 best stocks, and the Signal 0 means the most unsuccessful one. We could find that there is remarkable difference between return of Signal 0 and return of Signal 9. The returns we could get from the most successful portfolio are much higher than the poorest one. Also, we could calculate the Sharpe ratio to test our thoughts. As a result, the Sharpe ratio express an increasing trend from poorest to the best portfolio.

3. Alpha, Beta deliverables with CAPM model and FF3 model

1) CAPM model ***Alpha:***

As shown above, the ***'talpha'*** are the t values of different portfolio's alpha.

For every portfolio, the null hypothesis of alpha is that alpha=0. The table above shows that the ***talpha*** of all portfolios are between [-1.96,1.96]. Therefore, under 95% confidence level, CAPM cannot reject the null hypothesis. Alpha are 0, which means all these 9 portfolios do not have excess returns.

Beta:

As shown above, the ***'tbeta'*** are the t values of different portfolio's beta.

For every portfolio, the H0: beta=0. The table above shows that the ***tbeta*** of all portfolios are beyond [-1.96,1.96]. Therefore, under 95% confidence level, CAPM reject the null hypothesis. All beta in the table are more than 0, so all portfolios are positively related to the market. 2) FF3 model ***Alpha:***

The ***'talpha'*** in the above table are the t value of different portfolio's alpha.

For every portfolio, the ***'talpha'*** is between [-1.96,1.96. H0: alpha=0. Therefore, under 95% confidence level, FF3 cannot reject the null hypothesis. All alpha=0, which means all portfolios do not have excess returns.

Beta:

The ***'tbetamkt'*** in the above table are the t value of different portfolio's ***'betamkt'.***

For every portfolio, the ***tbetamkt*** is beyond [-1.96,1.96]. The null hypothesis is that ***betamkt***=0. All ***tbetamkt*** in the above table are more than 0. Therefore, under 95% confidence level, FF3 reject the null hypothesis of ***betamkt.*** The returns of the portfolios are positively related to the mkt.

The ***tbetahml*** in the above table are the t value of different portfolios' ***betahml.***

For some portfolios, the ***tbetahml*** are beyond[-1.96,1.96]. For other portfolios, the ***tbetahml*** are within (-1.96,1.96). So the return of a portfolio does not necessarily have a relationship with its ***hml.***

The ***tbetasmb*** in the above table are the t value of different portfolio's ***betasmb.***

For every portfolio, the ***tbetasmb*** is within (-1.96,1.96). The return of a portfolio is related to ***smb***. Some ***betasmb*** are positive, but some ***betasmb*** are negative. The return of a portfolio may be positively or negatively related to ***smb.***

4. Implement

The Hurst strategy we constructed is simple and straightforward. Thus, it has many flaws. There are several ways to implement the theory. We can mainly break down into three categories.

4.1 Mathematical method to construct the signal

The Hurst exponent estimator is important. The strategy we used computed the Hurst signal using R/S analysis, which explains relationship between the maximum range and the standard deviation of the stock returns in different fractal market. In this method, although robust, the Hurst exponent estimate is often biased. In testing, the Hurst exponents don't show much variance in short-term time series, and are close to 0.5 which are considered random walk. Other mathematical methods include more complex fractal motion theories. Options are Variance-time analysis or ***Detrended Fluctuation Analysis***, which are widely used in different areas.

4.2 Adjust Weight in portfolios

The Hurst exponent explained trading signals in stock level. In Jupyter, we simply use equal weight to demonstrate. To construct a signal level portfolio, we can adjust weight of different stocks. We can do empirical exercise using different weight, like value weight, risk-parity weight, to take a look at how the signal performs

4.3 Liquidity

Similarly, as the momentum strategy, the frequent rotation of the components of level portfolios requires a lot of trading. So the liquidity is also a major factor to concern. We demonstrate the trading volume figures in python, the illiquid stocks can exert great influence on portfolio performance. Also, frequent trading can result in great fees and taxes, which can bring down the actual excess return.

5. Cost & risk

Although the ***R/S*** analysis has been well developed and applied various new methods of to estimating Hurst exponent were proposed in the recent researches, such as variance-time analysis, adjusted rescaled range analysis, ***Detrended Fluctuation analysis*** and wavelet-based estimation. Therefore, one significant risk for Hurst exponent is that there is no standard method to estimate the Hurst exponent and the results from different methods varies. Each method has its own advantages, for example, in the case of a non-stationary time series represented, a trend and additive fractal noise, more accurate evaluation is obtained using this method while the adjusted rescaled range analysis is the most efficient with the lowest Mean Square Errors.

In addition, the signal depends on both the Hurst exponent and the beta. However, it is hard to determined the relative impact of each factors on signal. The Hurst index is between 0 to 1 while the value of beta varies in a much larger range. We only used the value of Hurst exponent and the sign of the beta in our strategy as we observed that the impact of value of beta is relatively small, but the value of the beta definitely carries the message theoretically.

Another risk of using this strategy are similar with the one attached with momentum investing. Focusing on the historical data would cause under-reaction of the new information and company status. Important event and announcements would significantly affect the stock price thus affect the stock return.

As the Hurst index depicts the long memory of a time series, it changes with times goes by. The bias depends on the true value of the degree self-similarity of the process and length of time series (Ludmila, Tamara & Zhanna, 2011). We found that the Hurst index is more indicative in shorter time series. In order to implement the strategy effectively, portfolios are reconstructed actively overtime. As the result, the management fee and the transaction cost will negatively impact the expected return of the investors.

Reference:

[1]Peters, E.: Chaos and Order in the Capital Markets – A New View of Cycles, Prices, and Market Volatility, John Wiley & Sons, Inc., New York, ISBN 978-0471533726, 1991

[2]Roel F. Ceballos1 & Fe F. Largo2: On The Estimation of the Hurst Exponent Using Adjusted Rescaled Range Analysis, Detrended Fluctuation Analysis and Variance Time Plot: A Case of Exponential Distribution, Imperial Journal of Interdisciplinary Research (IJIR) Vol-3, Issue-8, 2017 ISSN: 2454-1362

[3]Ludmila Kirichenko, Tamara Radivilova, Zhanna Deineko: Comparative Analysis for Estimating of the Hurst Exponent for Stationary and Nonstationary Time Series, International Journal "Information Technologies & Knowledge" Vol.5, 2011