Relative Value: Spread, Copula, Contra-Trend and XGB Approaches

King Wong

Abstract

I present an example of relative trading using three different approaches with a tech ETFs pairing with either its top holdings or stocks which are highly related to the sector. The backtest shows Copula > Spread > Contra-trend > in terms of mean Sharpe and alpha. The upside in capturing non-linear relationship using XGB to optimize the pairs trading is remained to be investigated further with more market pairs.

Introduction

Tasked with the exercise to construct a relative value strategy, I present a small example which can be used in a broader scale including other sectors inside or outside of the S&P500.

I have chosen the Technology Select Sector SPDR Fund(XLK) and a dozens of technology stocks as our toy example. I want to show the feasibility and the viability of three major approaches, 1. The Co-integration/ Spread approach 2. the copula approach 3. A contra-trend approach 4. a XGB approach. The former two have been widely discussed and implemented in literature (Liew, R., Wu, Y. Pairs trading: A copula approach. 2013) and I want to show how may trend statistics and ML add value in portfolio optimization/ market timing.

Methodology

Spread:

We can model the spread between the individual stock and the related etf as

log\_etf\_price - beta\*log\_stock\_price = v where v ~ N(mu, sig)

where beta is the log prices regression coefficient. The stationarity of the spread is ensured as the spread passed the ADF test for all pairs after normalize by its rolling mean and standard deviation. The strategy long the etf and short the stock when the spread is lower than -1.645 (stock overvalued, etf undervalued, spread narrows) and short the etf and long the stock when the spread is higher than 1.645 (stock undevalued, etf overvalued, spread widens). Other confidence interval can be chosen. For now, to keep things simple, long/ short same unit of stocks and etfs. Ideally we need to model the correct hedge ratio and bet size (1/2 Kelly’s Criteria)

Copulas:

As Rad, Hossein and Low, Rand Kwong Yew and Faff, Robert W., The Profitability of Pairs Trading Strategies: Distance, Cointegration, and Copula Methods (June 3, 2015) pointed out, normality assumption on the spread residuals may not always hold and it may miss out on trading opportunities. Copula function, a cumulative distribution describes the dependencies between two random uniform variables for each variable’s marginal distribution (Sklar’s theorem). We first transform the market returns of the pair into uniform distributed variables via their respective empirical cumulative distribution functions, which are modelled in-sample (temporarily ignore over-fitting for this example). Clayton’s copula is used here as there is an explicit function of the two uniform random variables defined by theta while other copula functions (e.g. Student’s t copula, Frank, Gumbel) should also be tested and decided based on AIC/ BIC criteria. The theta is obtained in-sample using kendall’s tau (see Genest, C., & MacKay, J. (1986). The Joy of Copulas: Bivariate Distributions with Uniform Marginals.). The theta, the two ECDFs ideally should be rolling or exponentially weighted as old returns distribution might not be representative of future returns. KL-divergence can reveal if the return distribution shifted significantly throughout time.

P(V≤v∣U≤u)=∂C(u,v)/∂u

P(U≤u∣V≤v) )=∂C(u,v)/∂v

Let u be the uniform random transformed price returns of the stock and v be the uniform random transformed price returns of the etf. We long stock and short etf if P(U≤u∣V≤v) < 0.05 and P(V≤v∣U≤u) > 0.95. Short stock and long etf if P(U≤u∣V≤v) > 0.95 and P(V≤v∣U≤u) < 0.05.

Quantile Portfolios:

We can form a quantile portfolio made up of the pairs which requires a continuous signal rather than a binary long/ short signal as above. We can simply use the normalised spread for the spread approach and the difference between the two conditional probabilities for the copula approach. The signal is ranked each day and the top 70% percentile are longed and the bottom 30% percentile are shorted. Other weighting schemes such as minimum-variance or maximised-Sharpe can also be used. Trading cost is ignored in this example, rebalancing as well, which ideally you want to incorporate in your full-backtest.

Contra-trend:

We can regress a short e.g. 5 day trend line to the spread in a rolling basis and obtain the t-value and p-value of the coefficient. We would then enter/ exit trade when there’s a reflection of trend/ no trend. Long stock and short etf when previous trend was positive. Short stock and long etf when previous trend is negative. Given both entries occurred when there’s no significant trend, p-value of the current trend is < 0.10.

XGB Portfolio:

We can train a classifier to decide which pairs to long and short. Since the trade entry and exit time varies (we enter one trade while close down the exact opposite trade), we need to label the data with the forward pair-trading return from each trading pairs calculated as price\_exit/ price\_entry – 1. This is done by taking the difference of the holding vector, select the ones that aren’t zeroes, calculate the percentage change and backshift t -1 to obtain the forward returns. The classifier is fitted to the sign of the forward returns along with all the features we have been using. Each row indexed by the date and the stock ticker.

First 5 years are being used as the training set and the rest as testing. The classifier generated signals are then equally weighted in the long and short positions to construct the final portfolio. We can also fit a XGB regressor and size our position in future development.

Backtest Results (2010 to 2020):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metrics\Portfolio | Spread | Copula | Contra-Trend | XGB(Contra-Trend Targeted)\* |
| Profitability(End Equity) | 5.164x | 6.489x | 2.278x | 0.988 |
| Alpha (daily) | 0.0006^ | 0.0007^ | 0.0003^ | -0.0000 |
| Beta (daily) | 0.0630^ | 0.0591^ | 0.0113 | -0.0004 |
| Volatility (daily) | 0.0100 | 0.0085 | 0.0089 | 0.0009 |
| Max Drawdown | -0.15 | -0.17 | -0.19 | -0.024 |
| Avg. Information Ratio (Annualised) | 0.254 | 0.530 | -1.18 | -50.11\*\* |

\* only 2015 to 2020 period is backtested.

\*\*extremely inflated due to low level of trading

^ statistically significant at 95% CI

We can see the traditional approaches still significantly outperform contra-trend or xgb-portfolio. But it’s too few data point to draw conclusion if XGB improves existing strategies. We also XGB model weights the conditional prob and trend statistics the highest in terms of feature importance.

Discussions and Potential Improvement:

1. Long/ short by hedge ratio, as stock and etf pairs trading returns are not 1 to 1 relationship
2. Bet-size should be gauged by the likelihood of mean-reversion(e.g. ½ Kelly), can be determined by the speed or momentum of the process
3. More stocks/etfs pairs to improve xgb trade selection training. The model would need more trade example in training set to determine entry and exit timing
4. More sector etfs to diversify the portfolio.