Optimal Parameters to Pairs Trading

Kirill Temlyakov

September 1, 2012

Abstract

Pairs trading is a very common trading strategy, and being able to obtain parameters that tell us when to trade and when to get out is of great importance. In this paper I propose a methodology that can improve the performance of traditional pairs trading strategy. I use stochastic beam search algorithm to find the best parameters in in-sample data within a given industry, and then test those parameters on out of sample data. My results outperform the results obtained by tools traditionally employed by the industry practitioners.

1 Work in progress

Please note that the following work is a work in progress and my results are preliminary. I intent to use more refined methods, for example explicitly include transaction costs. Moreover, I will include addition insights for example, I will compare my method performance with simple strategy for different industries.

2 Introduction

Pairs trading is a market neutral trading strategy that involves selecting two assets that move together, and trade when the gap between the two reaches a certain gap, a big point in this paper is how do we determine what this gap is. Profit is made when two assets comeback to the equilibrium. Even though pairs trading is common among practitioners the literature on it is scarce. Gatev and Goetzmann and Rouwenhorst (2006) [1] have investigated the performace of the methods employed by practitioners and found that pairs trading yields abnormally high return even if we account for transaction costs. Others have looked at what should be the criteria for choosing the pairs Vidyamurphy (2004) [2] in his book out lines an approach of selecting the pairs using cointegration, but does not test the approach empirically. In this paper, I adapt cointegration approach to finding the pairs. Elliott and van der Hoek and Malcolm (2005) [3] use Kalman filter to find the pairs by looking at parametric spread. In this paper I will focus on the choosing the parameters.

A very simple and common strategy goes as follows. Select a period, usually about a year long or about 250 trading days, find a pair of stocks that move together. Find the spread between the normalized prices of the stocks. Follow spread between the out of sample normalized prices, when the spread between the out of sample normalized prices exeeds two standard deviations in in sample spread, short the asset that has higher normalized price and buy asset with lower normalized price. Now you hold this position. If the differences between normalized prices goes back to zero, you sell the asset you hold, and buy the asset you shorted, thus making profit. Otherwise you just close out on the position at the end of the specified period, usually 6 or 12 months.

This methodology, though, having a lot experience is chosen purely ad hoc, and might not work best in some situations. I propose a methodology that selects a number of pairs from each industry group. Then finds best (or approximately best) parameters for the returns next period. I can do that since I know the returns. And then applies those parameters for the entire industry group on the next sample. The parameters do vary between industries.

I propose one additional parameter, in addition to *when to trade* and *when to trade back*, I add *loss control*. *Loss control* tells us when to close out. Let's say that two price paths have drifted so far apart that it is more likely that they will keep going further apart, than that they will come back. In this case it makes sense to close out the position. And *Loss control* tells us exactly what that parameter will be. In the previous simple example, *when to trade* is sum of two standard deviation of the normalized prices last period, *when to trade back* is 0, and*loss control* is infinity.

The crucial part of my methodology relies on the ability to find parameters that are close to optimal when we know the returns for the next period. Return in the subsequent period is a function of the three parameters. The difficulty lies in the fact that the terrain of the function has lots of local maximums, and thus any hill climbing algorithms do not work well. Grid search is extremely inefficient. Therefore, I use stochastic beam search, its ability to jump from place to place ensures that it will not get stuck on a local maximum. Stochastic beam search is a heuristic algorithm; it is working to improve the solution based on past experience. Stochastic beam search generates a population of parameters chosen randomly, removes n weakest performing elements, from the remaining population selects n elements that are modified and added back to the population. The probability of being selected for reproduction is often proportional to strength on the element. Stochastic beam search does not guarantee that I will find the global maximum, but I should come pretty close. The exact details on construction of the stochastic beam search algorithm employed in this paper is in the Methodology section.

Stochastic beam search's close relative, another heuristic algorithm, genetic algorithm is a frequent guest in the literature. Franklin and Karjalainen(1999) [4] use genetic algorithm and genetic programming to establish rules for technical trading. Cochran and Lertwachara (2007) [5] use genetic algoritm and neural networks to estimate analyst recommendations. Kapoor and Dey and Khurana (2011) [6] propose a genetic algorithm to be used for technical trading systems.

The main difference between stochastic beam search and genetic algorithm is that genetic algorithm uses two methods to update the population, mutation and crossover. The way population is updated in stochastic beam search is basically mutation; however, stochastic beam search in completely missing the crossover part. Crossover results when we have two parents, and a child inherits some genes from each parent, and other genes could be chosen at random. In our case we have three genes, the parameters: *when to trade*, *when to trade back*, and *loss control*.

I chose stochastic beam search over genetic algorithm, because I felt that crossover operation does not provide enough benefit in this case.

3 Data

All data for this paper is collected from CRSP. I used, along with dates and permnos, holding period return and HSICIG, which identifies the industries. I am using daily data for time period between January 01, 1993 and December

31, 2011. The returns are not adjusted for risk free rate, and through out this paper I do not adjust for risk free rate.

4 Methodology

In this section I will outline in detail the methodology I used in this paper. First I will explain in detail the stochastic beam search, second I will explain criteria used in finding the pairs, then I will explain the trading strategy, and finally I will put everything together and explain how the paper is constructed.

4.1 Stochastic Beam Search

I use the stochastic beam search algorithm as outlined by Norvig and Russell (2003) [7].

I select n elements of the population. Each element is chosen from normal distribution and absolute value of the term is taken.

when to trade N(0,1)

when to trade back N(0,2)

loss control N(0,5)

Therefore, I am imposing some structure; however, it is really easy for the algorithm to move around.

After that I conduct m iterations, at each iteration I remove k weakest elements, and replace them with mutated k strongest elements. The measure of strength is performance of the portfolio next period. For mutation, I take an element, keep 2 of its parameters, and change the third. The parameter that is being changed is selected randomly. I change it by adding a value taken from *N(0,2)*.

Throughout the paper my I use: $n = 40$, $m = 10$, $k = 5$.

And we pick parameters that yield the highest return.

4.2 Normalized Prices

I construct normalized prices by using holding period return from the CRSP database, using this formula:

$$pn_{t+1} = pn_t * (1 + r_{t+1}) \tag{1}$$

I set first value of the normalized prices to 1 by default.

4.3 Finding The Pairs

I consider a pair to be fit for pairs trading if the series in the pair are cointegrated. I check this by regressing one variable on the other and applying augmented dickey-fuller test to the residual. If the test statistic is lower then 5 percent rejection value. I consider a pair to be cointegrated.

4.4 Trading Strategy

Given the way we calculate normalized prices both assets start with normalized price of 1, and then depending on the returns the normalized price changes. When the difference between normalized prices of assets is equal to or greater than *when to trade* parameter, we buy the asset that has a lower normalized price and we short the asset with higher normalized price. We always buy the same value worth as we short. If we did trade, we can not trade anymore. Suppose we made a trade, if the difference between normalized prices becomes less or equal to *when to trade back*, we sell the asset we bought, and buy the asset we shorted. The game starts over and we can keep doing this until the end of the period, which in my case is 250 trading days. If however, the difference reaches *loss control*, we close out the position and get out. At the end of the trading period, we automatically close out the positions.

4.5 Construction of the Paper

From the CRSP data I obtained the data from January 01, 1993 to December 31, 2011. I look at the first period in the data, it is 750 values; I am looking only at assets that have all the observations. I divide the data into groups by industries; I kept only industries that have more than 100,000 observations. This left me with 76 industries. I estimate different parameters for every industry group; I do the following for every industry: I divide the 750 observations into 3 sub-periods: 0:250, 250:500, 500:750. We can think of 0:500 part as 500 trading days before today, and 500:750 as 250 trading days after today; the last 250 observations is the part we do not know.

1)I use the first sub-period to find all cointegrated pairs, if I have more then 10, I take 10 cointegrated pairs with lowest augmented dickey-fuller test statistic.

2)I know the pairs in 0:250 and I know the returns in 250:500. I use stochastic greed search to find the best parameters.

3)I find cointegrateded pairs in 250:500.

4)Use the the parameters found in 2) on and the pairs found in 3) to trade in period 500:750. And record the profit for each industry. If industry does not have any cointegrated pairs for the given time sample zero profit is recorded.

Overall in our data we have 17 periods containing 750 such observation; hence, we repeat the above process 17 times, and record the overall profit. The results of the profit are displayed in Table 1.

It is important to note that stochastic beam search is stochastic, and therefore the results given by the algorithm will vary from test to test. For this reason we are running the test two times.

4.6 Trading Strategy for Simple Method

I do not need the data in period 0:250. I only use data in period 250:500 to find the cointegrated pairs in each industry. I apply the same methodology here as in the above trading strategy. I also use the same methodology to construct normalized prices. Once I identified the cointegrated pairs. I trade those pairs in the future period 500:750. If sum of standard deviations of each pair in period 250:500 we make a trade; we buy the asset with lower normalized price and short the asset with higher normalized price. Once we made the trade we do not make anymore trades until either difference between normalized prices is zero and we trade back, sell the asset we bought and buy the asset we shorted, and start over or we reach the end of period and have to close out our position.

4.7 Calculating the Return

I calculated the return as follows:

$$return = sumprofit/(numberoftotalassets) * amountinvested * 100percent \tag{2}$$

For example, if I decided to short 1000 dollors worth of one asset and buy 1000 dollars worth of the other asset per trade and made 500 dollar profit at the end. My return would be 500/2000 = 25 percent.

5 Empirical Results and Conclusion

Table 1 shows the results of my proposed methodology under two different tests, and the results for the simple strategy. The mean return for my methodology in test one is 14.34 percent (the returns are not adjusted for risk-free rate) and for test two 14.04 percent. The outcome given by the method does not vary much for any of the periods. The method performs better than the simple strategy which yields a return of 11.66 percent. However, most advantage can be attributed to the early periods. Moreover we are not accounting for bid-ask spread and other transaction costs, which can be substantial for trading strategy that relies on fairly frequent trading. A further study that takes into account those aspects is of great importance.

References

[1] E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst, "Pairs trading: Performance of a relative-value arbitrage rule," *The Review of Financial Studies*, vol. 19, no. 3, pp. 797–827.

[2] V. G., *Pairs Trading Quantitative Methods and Analysis.* John Wiley and Sons, Canada., 2004.

[3] R. Elliott, J. van der Hoek, and W. Malcolm, "Pairs trading," *Quantitative Finance*, vol. 5, no. 3, pp. 271–276.

[4] F. Allen and R. Karjalainen, "Using genetic algorithms to find technical trading rules," *Journal of Financial Economics*, vol. 51, pp. 245–271.

[5] J. Cochran and K. Lertwachara, "Selecting stocks using a genetic algorithm: a case of real estate investment trusts," *International Journal of The Computer, the Internet and Management*, vol. 15, no. 2, pp. 20–31.

[6] V. Kapoor, S. Dey, and A. Khurana, "Genetic algorithm: An application to technical trading system design," *International Journal of Computer Applications*, vol. 36, no. 5, pp. 44–50.

[7] S. Russel and P. Norvig, *Artificial Intelligence A Modern Approach.* Prentice Hall, 2003.

Table 1: Displays the returns for each period. In the first two columns are results from my proposed methodology conducting two different tests. In the last column are results from the simple strategy. The mean returns are displayed at the bottom of the table.

Period	Return test 1	Return test 2	Simple Strategy Return
1	28.88	30.56	16.93
2	22.40	23.24	14.76
3	21.97	21.07	13.25
4	24.15	23.57	20.56
5	17.58	16.72	8.63
6	21.58	21.39	24.20
7	24.83	22.69	22.06
8	19.73	19.14	17.76
9	7.34	6.74	-0.31
10	3.46	2.95	5.42
11	2.51	2.36	0.32
12	3.50	3.80	2.68
13	0.18	0.39	-0.04
14	8.73	10.00	16.88
15	22.86	21.2	23.00
16	7.81	6.51	5.85
17	6.21	6.28	6.25
Mean Return	14.34	14.04	11.66