An Evaluation of Pairs Trading in Commodity Futures Markets

Li, Shen
China Asset Management Co., Ltd.
Beijing, 100033, China
shenli19941220@163.com

Kun, Shen
China Asset Management Co., Ltd.
Beijing, 100033, China
kongk@chinaamc.com

Chao, Yi
China Asset Management Co., Ltd.
Beijing, 100033, China
yic@chinaamc.com

Yixin, Chen
China Asset Management Co., Ltd.
Beijing, 100033, China
chenyx@chinaamc.com

***Abstract*—Pairs trading is an algorithm to trade two cointegrated assets in order to get neutral profits theoretically. This paper is aimed to demonstrate the cointegration concepts and pairs trading algorithms. Also, we need to take reality into account, especially margins and commission fees. We use grid search methods to find ideal combinations of trading signals. It turns out that this trading algorithm can achieve great performances.**

Keywords—cointegration, pairs trading, commodity futures

I. INTRODUCTION

Pairs trading is a market-neutral strategy. We need to find two assets whose prices are cointegrated. When it is believed that one asset will outperform the other, then we will long (buy) the former and short (sell) the latter. As a result, we can achieve risk-neutral profits for a good chance when these two asset prices moving to the middle point and we close our trading positions in the end. How can we find these two cointegrated assets and trade them in some ratios? And when should we close our positions in these two assets? We will elaborate these in our paper.

In section Ⅱ, we show some basic concepts, such as stationarity and cointegration. In section Ⅲ, we demonstrate the trading algorithms. In section Ⅳ, we show settlement algorithms in Chinese futures exchanges. In section Ⅴ, we try grid search methods to find the ideal combination of coefficients. In section Ⅵ, we contrast our work with referred work and we want to look forward to future improvements.

II. BASIC CONCEPTS

A. Stationarity

Definition 1: Time series $\{X_t\}$ is strictly stationary if $F(x_1, \dots, x_t) = F(x_{1+\tau}, \dots, x_{t+\tau})$ for any τ, t. Therefore, the mean and variance of $\{X_t\}$ are both constants irrelevant of time, and covariance $Cov(x_t, x_{t+\tau})$ depends only on τ irrelevant of time, too.

B. Integration and Order

Definition 2: Time series $\{X_t\}$ is integrated of order 1, I(1) if $\{\Delta X_t\}$ is stationary, where $\Delta x_t = x_t - x_{t-1}$. Most economic time-series data are I(1).

C. Cointegration

Definition 3: When time series $\{X_t\}$ and $\{Y_t\}$ are cointegrated, we have

$$Y_t = \alpha + \beta X_t + \epsilon_t, \text{ where } \epsilon_t \sim I(0) \tag{1}$$

Granger has proved that if two time-series vectors are cointegrated, then deviations from long-run equilibrium are stationary, even if the series themselves are nonstationary [2][3]. So we can judge whether two assets are cointegrated or not based on (1).

III. PAIRS TRADING

Pairs trading is an algorithm involving two assets with a significant cointegrated relationship. We should open opposite positions in these two assets when their prices deviate from each other in the short run, and close all positions when their prices return to the long-run equilibrium, so that we can achieve profits during the process.

A. Dicky-Fuller Test

The first step is to find two cointegrated assets based on (1). And Dicky-Fuller test is a more formal method developed from (1). In DF test, we have two asset price series, $\{A_t\}$ and $\{B_t\}$, where $y_t = log\, A_t - log\, B_t$, and we regress Δy_t on y_{t-1}

$$\Delta y_t = y_t - y_{t-1} = \mu + \gamma y_{t-1} + \epsilon_t \tag{2}$$

$$H_0: \gamma = 0, H_1: \gamma < 0 \tag{3}$$

If p-value of regression coefficient γ is under the critical point, then $\{y_t\}$ is stationary, because $y_t = (1+\gamma) * y_{t-1} + \mu + \epsilon_t$ and $1+\gamma$ is less than 1. Therefore, two asset price series, $\{A_t\}$ and $\{B_t\}$, are cointegrated.

B. Financial Background Knowledge

It is appropriate to use commodities in this trading model which involves long and short trading positions. In reality, commodities are traded in futures contracts instead of real commodities. Futures are financial contracts that obligate the parties to transact an asset at a predetermined future date and price. Long position means that we enter into contracts to buy

IEEE 978-1-7281-6251-5/20/$31.00 ©2020 IEEE

commodities in the future, and short position means that we enter into contracts to sell commodities in the future. When we close our positions in commodities, we enter into the opposite futures contracts. And we focus on dominant-continuous futures contracts, which are traded continuously and in the largest volume. Therefore, we use the word ‘futures’ instead of ‘commodities’ in the following context.

We trade futures on margins, which means that we only pay fixed percentages of futures contract values, called margins, instead of the total value. Therefore, we expand profits and losses at the same time.

C. *Pick Cointegrated Futures*

Before we could find any relationships, be careful that we should not use future data. Therefore, after we test cointegration relationships in history data, this part of data won’t be used on trading any more, as shown in Fig. 1.

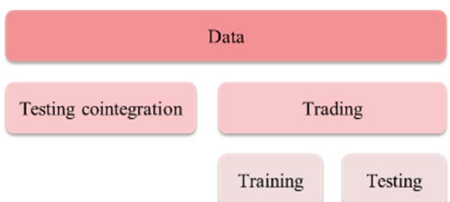

Fig. 1. Data usage

We take out data before December 31st, 2015 to perform cointegration and correlation tests. Table I shows part of Pearson correlation coefficients and Dick-Fuller test results among futures.

TABLE I. CORRELATION AND COINTEGRATION TEST RESULTS

Futures 1	Futures 2	Correlation Coefficient	Correlation P-value	Cointegration Score	Cointegration P-value
RB.SHF	WR.SHF	0.98346	0.00000	-6.41821	0.00000
CF.CZC	PP.DCE	0.79336	0.00000	-5.64534	0.00001
PB.SHF	Y.DCE	0.89595	0.00000	-5.12627	0.00010
JD.DCE	PP.DCE	0.87847	0.00000	-4.87012	0.00028
PB.SHF	RU.SHF	0.92734	0.00000	-4.66270	0.00065

Here we choose CF.CZC (cotton) as f1 and PP.DCE (polypropylene) as f2, both of which are highly cointegrated, although their correlation is not the strongest. And settlement prices after December 31st 2015 are plotted in Fig. 2. This is an exciting result because we didn’t use data after December 31st 2015 to test cointegration, but these two futures follow perfect cointegrated relationship anyway.

Fig. 2. Cotton and polypropylene settlement prices

D. *Trading Algorithms*

After we pick two statistically significant cointegrated futures, we build a linear regression model in (1) with settlement prices $\{f1_t\}$ and $\{f2_t\}$, where β, as the relative price ratio between f1 and f2, would follow mean-reversion trend the long run. However, β would fluctuate in the short run. Therefore, we allow β to fluctuate in a fixed range. When β hits the upper bound, it means that $f1_t$ is expensive and $f2_t$ is cheap, so that we short 1 share of $f1$ and long β shares of $f2$. When β hits the lower bound, we trade in the opposite direction. And when β returns to intervals between the upper and lower bound, we close our positions.

β is the relative price ratio between two futures, so it varies with different pairs of futures. We normalize β into z score

$$z\ score = \frac{\beta - mean(\beta)}{std(\beta)} \sim N(0,1) \tag{4}$$

which applies to all contracts and won’t vary with different pairs.

When we trade continuously, we have to use a rolling index instead of a once-for-all index. So, we should calculate rolling z score everyday

$$z\ score_{roll} = \frac{MV_5(Ratios) - MV_{60}(Ratios)}{std_{60}(Ratios)} \tag{5}$$

Here, we use 5 trading days as a short-term trend and 60 trading days as a long-term trend. As Fig. 3 shows, rolling z-score is more stable without trivial fluctuations. When $z\ score_{roll}$ hits 1, we short f1 and long f2; when $z\ score_{roll}$ hits -1, we short f2 and long f1; and when $z\ score_{roll}$ returns into [-0.5, 0.5], we close our positions. Here, 5 days as short-term period, 60 days as long-term period, ± 1 as outer bound and ± 0.5 as inner bound, can all be fixed by grid search methods.

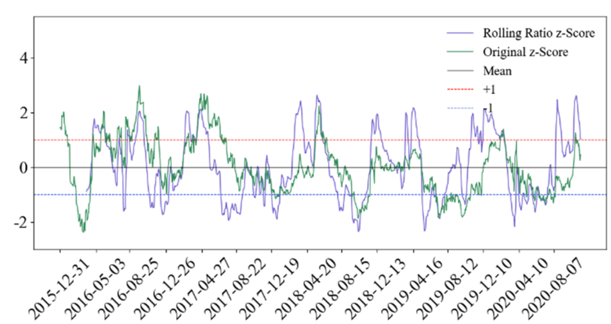

Fig. 3. Z-score and rolling z-score plots

After we find trading signals, they are projected back onto futures contracts plots, as showed in Fig. 4.

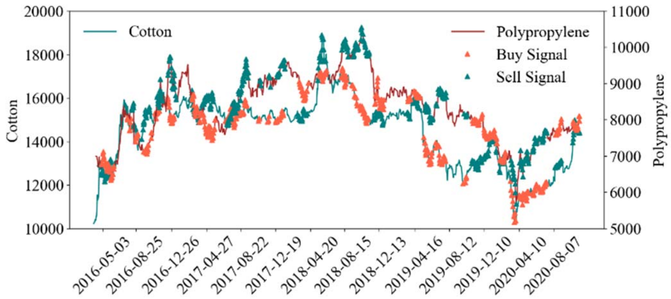

Fig. 4. Trading signals in contracts

We can start to trade everyday based on trading signals from Fig. 4. Here we simplify the whole process, assuming that we can trade at settlement prices. Price ratios won't always be an integer, so we open 10 shares of CF.CZC (f1) and get a rounding result of ratios*10 shares of PP.DCE (f2) each time when signals are sent out. And in this simplified trading model, we short one part and then use the money to long the other. So initial investment is zero. As we can see in Fig. 5, total profits are really awful.

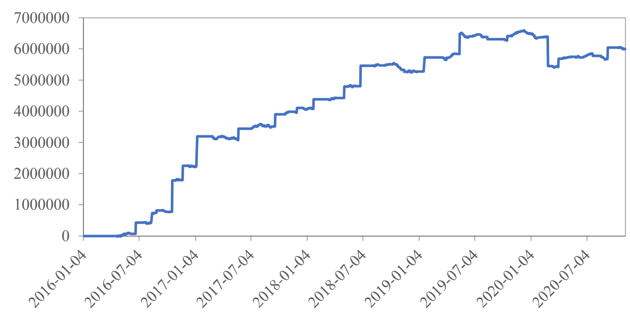

Fig. 5. Total accumulated returns without margins

IV. Settlement System

A. Details

As we move forward, we need to consider settlement algorithms in Chinese futures exchanges. And in this paper, we trade futures on daily basis, which means that we don't have to consider intraday trading profits and losses.

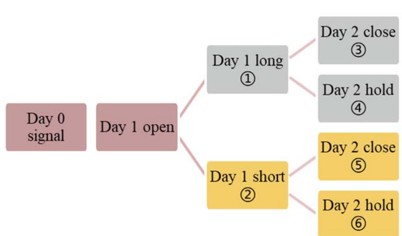

Fig. 6. Daily trading process

As Fig. 6 shows, we get trading signals at the end of day 0 with settlement prices. And we have put m_0 money in our settlement account on day 0. We open positions at the beginning of day 1.

Situation ①: If we open positions by longing $share_1$ at the price of $open_1$ at the start of day 1 and get settlement price $settle_1$ at the end of day 1, then we have opening-profits-losses (as openPL, the same below) in (6), margins in (7) and money in the settlement account in (8)

$$openPL = (settle_1 - open_1) * share_1 \tag{6}$$

$$delta\ margin_1 = settle_1 * share_1 * margin\ rate \tag{7}$$

$$m_1 = m_0 + openPL - delta\ margin_1 - commission_1 \tag{8}$$

And we have to compare m_1 with total margin in account to decide whether we should receive a margin call. If $m_1 < delta\ margin_1$, $delta\ margin_1 - m_1$ should be put into account to cover the gap. Therefore, total money at the end of day 1 is

$$m_1 + max(delta\ margin_1 - m_1, 0) \tag{9}$$

Situation ③: following ①, we close our day 1 positions by shorting $share_1$ at the price $close_2$ at the end of day 2. Then we have closing-history-position profits and losses (as closePL, the same below) in (10), total money in account in (11)

$$closePL = (close_2 - settle_1) * share_1 \tag{10}$$

and total money in our account is

$$m_1 + closePL + total\ margin_1 - commission_2 \tag{11}$$

Situation ④: following ①, we keep holding our day 1 positions and the settlement price is $settle_2$ at the end of day 2. Then we have holding-history-position profits and losses (as holdPL, the same below) in (12)

$$holdPL = (settle_2 - settle_1) * share_1 \tag{12}$$

and total money in our account is

$$m_2 = m_1 + holdPL + total\ margin_1 \tag{13}$$

And we have to compare m_2 with $total\ margin_2 = settle_2 * share1 * margin\ rate$. Therefore, total money in account is

$$m_2 + max(m_2 - total\ margin_2, 0) \tag{14}$$

Situation ②: If we open positions by shorting $share_1$, then we have opening-profits-losses in (15)

$$openPL = (open_1 - settle_1) * share_1 \tag{15}$$

Situation ⑤: following ②, we close our day 1 positions by long $share_1$ at the price $close_2$ at the end of day 2,

$$closePL = (settle_1 - close_2) * share_1 \quad (16)$$

Situation ⑥: following ②, we keep holding our day 1 positions and the settlement price is $settle_2$ at the end of day 2,

$$holdPL = (settle_1 - settle_2) * share_1 \quad (17)$$

At day 2, besides closing or holding the original contracts, we can also open new positions. And all is the same as this whole process.

B. Simulations

So, after we consider margins, we have new results as showed in Fig. 7 and Fig. 8.

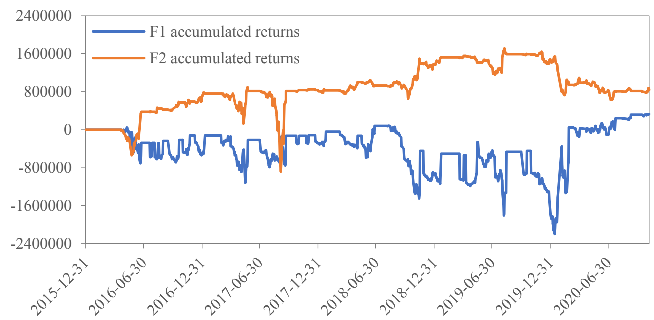

Fig. 7. Individual accumulated profits

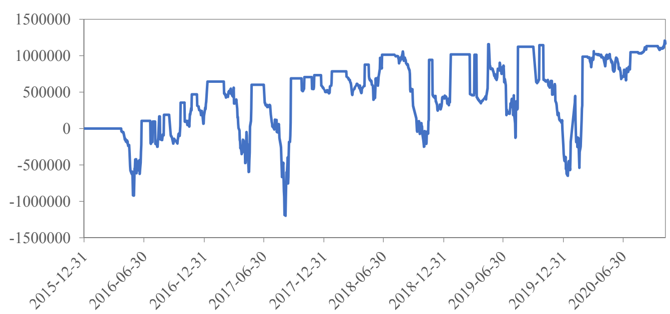

Fig. 8. Total accumulated returns with margins

As we can see, there exit some significant drawdowns, as margin trades will expand losses and profits at the same time. And the total accumulated returns with margins are just a quarter of those in Fig. 5. Meanwhile, cotton and polypropylene are cointegrated during this whole history. So, when we long one and short the other, these two accumulated return plots are in opposite directions as shown in Fig. 9. Be that as it may, we believe that short-run deviations will vanish and we can achieve profits during this restoration process.

V. GRID SEARCH

A. Train and Test

In section Ⅲ and Ⅳ, we used predetermined coefficients, such as upper bound, lower bound, inner bound, short-term period and long-term period, all of which could be determined by grid search methods. Here, we try to find out the best combination of coefficients both in training and testing data.

First of all, we still use data after December 31st 2015 and split this part data into training and testing samples, accounting for 60% and 40% respectively.

Then, for different coefficients, we have different candidates. Short-term periods are chosen among [5, 6, 7, 8, 9, 10, 15, 20, 25, 30] days, long-term periods are chosen among [50, 60, 70, 80, 90, 100, 120, 250] days, upper bound are chosen among [1, 1.5, 2, 2.5], lower bound are chosen among [-2.5, -2, -1.5, -1], and inner bound are chosen among [0.1, 0.2, 0.3, 0.4, 0.5]. Therefore, there are 6400 combinations of coefficients. Here, we can calculate accumulated returns of each coefficient combination both in training and testing samples. And we sort accumulated returns from the biggest to the smallest in training and testing samples respectively. We sum up training ranks and testing ranks, and then sort this 'sum rank' from the first to the end. As we can see in TABLE Ⅱ, the best combination of coefficients in training data doesn't necessarily conform with that in testing data.

TABLE II. BEST 10 RESULTS

Short Window	Long Window	Upper Bound	Lower Bound	Inner Bound	Training Rank	Testing Rank
20	250	1	-1	0.2	6	12
30	250	1	-1	0.1	7	62
25	250	1	-1	0.1	4	84
20	250	1.5	-1	0.2	35	53
20	250	1	-1	0.1	1	124
15	250	1.5	-1	0.1	3	129
5	80	1	-1	0.1	64	76
20	120	1	-1	0.5	81	69
20	250	1.5	-1	0.1	5	165
7	120	1.5	-1	0.1	71	100

TABLE III. WORST 10 RESULTS

Short window	long window	Upper bound	Lower bound	Inner bound	Training rank	Testing rank
20	70	1	-1	0.5	6370	5862
15	60	1	-1	0.1	6368	5879
25	80	1	-1.5	0.4	6338	5911
25	80	1	-2	0.4	6339	5912
25	80	1	-2.5	0.4	6340	5913
20	70	1	-1	0.1	6392	5946
25	80	1	-1	0.4	6400	5949
15	70	1	-1.5	0.1	6321	6051
15	60	1	-1	0.2	6399	6019
15	70	1	-1	0.1	6371	6339

B. Analysis of Coefficients

Then we want to see what's going on with the distribution of these coefficients. We plot the histograms of the best 200 and worst 200 results.

In Fig. 9 and Fig. 10, x-axis displays rank numbers. Best conditions are those where long windows are around 120 days and short windows are around 10 days. In contrast, worst conditions are those where long windows are 70 days and short windows are around 20 days.

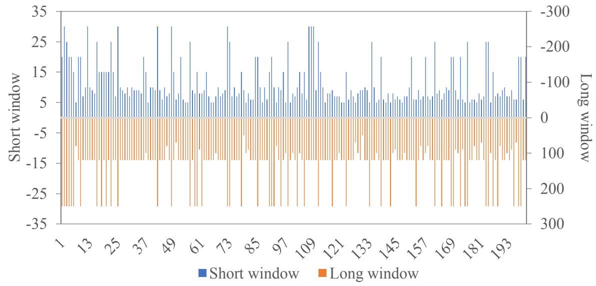

Fig. 9. Best 200 short-long rolling window combinations

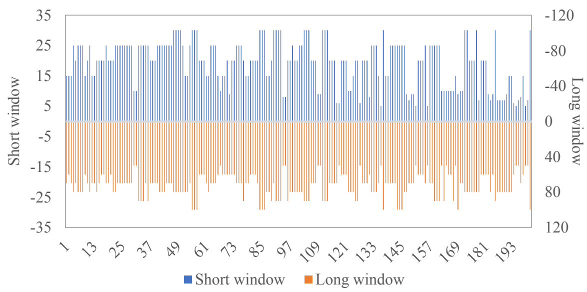

Fig. 10. Worst 200 short-long rolling window combinations

In Fig. 11 and Fig. 12, x-axis shows rank numbers. The best results show that most lower bounds are located in -1, while the worst results show that most lower bounds are located at -2.

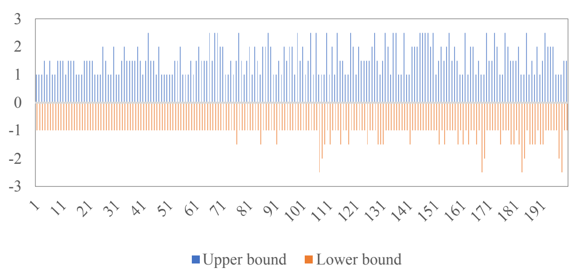

Fig. 11. Best 200 upper and lower bound combinations

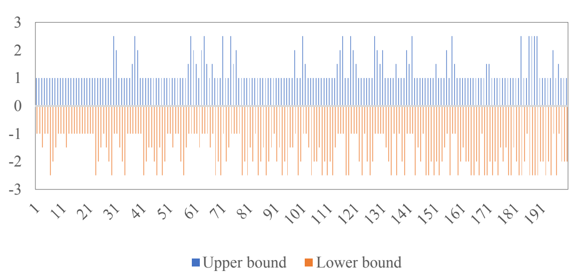

Fig. 12. Worst 200 upper and lower bound combinations

In Fig. 13 and Fig. 14, inner bounds are irrelevant because there are not outstanding differences.

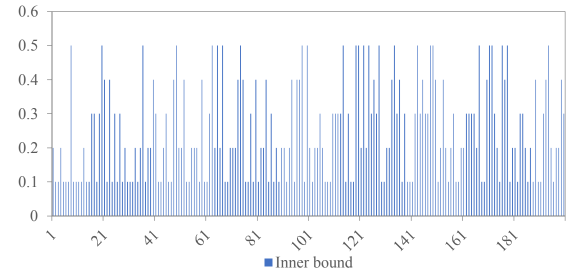

Fig. 13. Best 200 inner bounds

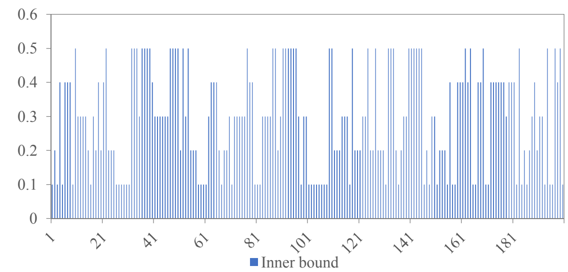

Fig. 14. Worst 200 inner bounds

So, conclusions are drawn that we should keep short rolling window in 10 days, long rolling window around 120 days, upper bounds in 1 and lower bounds in -1, which will produce ideal results.

C. Best 5 and Worst 5 Accumulated Returns Plots

As is showed in Fig. 15 and Fig. 16, best and worst groups share much more similar trends in figures respectively. Not only is the drawdown much smaller in the best groups, but also trends are more stable in the best groups.

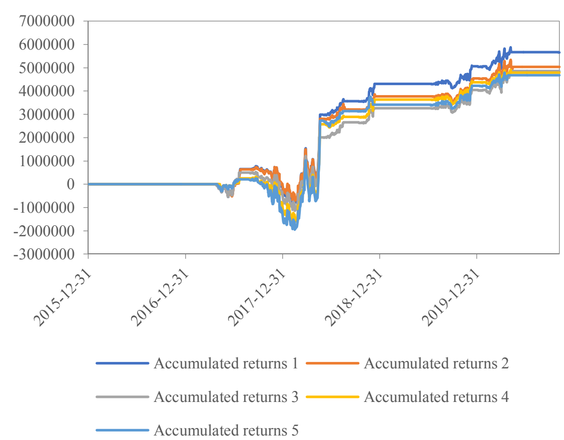

Fig. 15. Best 5 groups

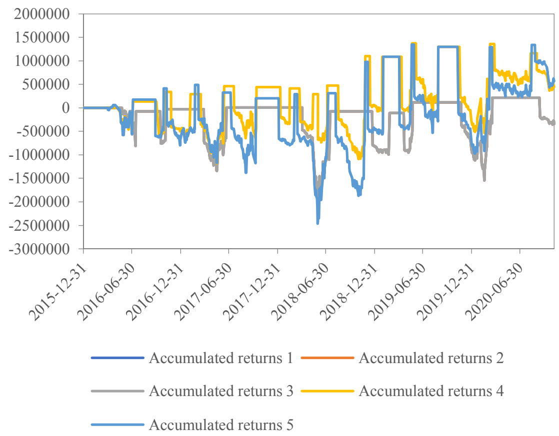

Fig. 16. Worst 5 groups

VI. Summary and Future Improvement

A. Comparison

We referred to a paper of Daniel Herlemont [1]. In his paper, he explained basic concepts in details. And more importantly, he talked about risk control ideas and trading algorithms. But there weren't so many examples. He mainly focused on stocks pairs trading instead of commodity futures. In our paper, we showed specific examples and tried grid search to find the optimal combinations of coefficients. So that is main improvements in our work.

B. Future improvements

We could actually add minute trading in this algorithm. And in intraday trading, we have to label every position so that we can trace them easily and build up a more flexible settlement system. Also, beside rolling-score windows, we could take more technical information into account, such as price-volume ratios and macro information, in order to judge whether or not we should enter into a pairs trading.

References

[1] Herlemont, Daniel. "Pairs Trading." Convergence Trading, Cointegration, URL http://www. yats. com/doc/cointegration-en. pdf (2004).

[2] Granger, C. W. J. . "Some properties of time series data and their use in econometric model specification." Journal of Econometrics 16.1(1981):121-130.

[3] Granger, Robert F. Englec. W. J. . "Co-Integration and Error Correction: Representation, Estimation, and Testing." Econometrica 55.2(1987):251-276.