

SELECTING STOCK PAIRS FOR PAIRS TRADING WHILE INCORPORATING LEAD-LAG RELATIONSHIP

Authors: Kartikay Gupta (corresponding author), Niladri Chatterjee

Emails: maz158144@maths.iitd.ac.in, niladri@maths.iitd.ac.in

Mathematics Department, IIT Delhi, New Delhi, Pin-110016, India

Keywords: correlation, distance metric, pairs trading, finance, lead-lag relationship

Abstract

Pairs Trading is carried out in the financial market to earn huge profits from known equilibrium relation between pairs of stock. In financial markets, seldom it is seen that stock pairs are correlated at particular lead or lag. This lead-lag relationship has been empirically studied in various financial markets. Earlier research works have suggested various measures for identifying the best pairs for pairs trading, but they do not consider this lead-lag effect. The present study proposes a new distance measure which incorporates the lead-lag relationship between the stocks while selecting the best pairs for pairs trading. Further, the lead-lag value between the stocks is allowed to vary continuously over time. The proposed measures importance has been show-cased through experimentation on two different datasets, one corresponding to Indian companies and another corresponding to American companies. When the proposed measure is clubbed with SSD measure, i.e., when pairs are identified through optimising both these measures, then the selected pairs consistently generate the best profit, as compared to all other measures. Finally, possible generalisation and extension of the proposed distance measure have been discussed.

1. Introduction

Pairs Trading [1] is done in the financial market to earn money from the uncertain movements of the stock prices. Basically, a trading strategy exploits a relationship between two stocks to earn money. This relationship can exist and can be exploited in many different ways. Here, a stock is bought, and simultaneously the other stock is shorted, whenever the equilibrium relationship between the two stocks gets temporarily disturbed. A profit is realised by squaring off the positions whenever the stocks fall back in the equilibrium relationship.

Pairs trading has been used by financial engineers for decades. As mentioned in the article [2], Alfred Winslow Jones (around 1950) used the concept of pairs trading for going long on certain stocks while short on others. Though, several authors [3]–[5] are of the view that pairs trading does not lead to any significant profits when done over a long period. Still, it is extensively used in today's financial market for earning money [6]. A recent survey on different pairs trading strategy is [1].

The measures commonly employed to select pairs for pairs trading are correlation and SSD [7], which are described later. It is often observed that certain stocks are highly correlated to each other at some 'lead or lag'. The above measures do not take into account this phenomenon of 'lead-lag' relationship. This phenomenon is important and can not be ignored while identifying stock pairs.

In this paper, we present new distance metrics for pairs selection, which overcomes this shortcoming. We also showcase the effectiveness of another distance measure CCT [8], which has not been earlier used for this task. The significance of these new measures has

been validated through experimentation, as described in this paper. Traders would undoubtedly find it beneficial to analyse the values of these new measures, along with other technical indicators, while picking up stocks for pairs trading.

The rest of the paper is organised as follows. Section 2 gives a literature review. Section 3 elaborates proposed distance measures. Section 4 presents the experiments and results. Section 5 discusses various aspects of this research work. Section 6 concludes.

2. Literature Review

The present paper proposes a new distance metric for pairs selection. The proposed distance measure has been inspired by the CCT measure proposed in [8]. We also showcase CCT measure effectiveness in pairs selection. It is sometimes observed that some stock pairs may not have high Pearson correlation coefficient between them, but they are highly correlated at certain lead-lag. This phenomenon is generally referred to as ‘lead-lag effect’. As reported in [9], ‘lead-lag effect’ the lagger (or follower) stock’s price tends to partially reproduce the oscillations of the leader (or driver) stock’s price, with some temporal delay. This lead-lag effect has been studied and analysed in financial econometrics. High-Frequency Data in the financial market is gathered at irregular intervals, which makes it challenging to decipher the lead-lag relationship between two different stocks/markets. Thus, an estimator is proposed in [10] which better estimated the cross-covariance by avoiding imputation and using all available transaction. The lead-lag relationship in KOSP 1200 spot market, it’s future market and it’s option market is empirically examined and commented upon in the study done by Lee et al [11]. In the study conducted by Robert et al [12], certain properties of the covariance matrix of increments of two Gaussian processes, partially correlated at some time $-\text{lag}$, is studied.

Another distance measure of significance is the Dynamic Time Warping (DTW) measure

[13]. DTW is generally considered as the best distance measure for time series mining tasks across virtually all domains [14]. DTW measure is especially of advantage in speech recognition [15] where it can decipher the sounds of different words, even when different parts of the word have different elongations. Zhu et al. [16] tried to reduce the time complexity of the DTW measure by approximating its value. In work by Silva et al [14] the effects of relaxing various constraints on the DTW distance measures are studied.

As discussed earlier, stock prices tend to be correlated at some lead or lag. Also, the lead-lag value (or the temporal difference) between the prices of stock pairs may change over time. The DCCT measure allows the lead-lag relationship to vary continuously over time.

It subtly combines the properties of Dynamic Time Warping (DTW) measure with the CCT measure [8].

The following is a brief description of two distance metrics, which in the past have been used for pairs selection. These metrics will be used for comparison with the proposed metric and the CCT metric.

1. Pearson Correlation

Pearson correlation is frequently used as a metric for pairs selection. The stocks which have higher Pearson correlation, their prices tend to move more synchronously together. The work by Wang et al [17] analyses different correlation measures w.r.t. pairs trading.

It is reported in [18], that the different correlation measures do show important differences in terms of return and risk.

2. SSD measure

The SSD measure is defined as the sum of squared differences of normalised price series of the two stocks. In the present experiments, the time series $\{X_1, X_2, \dots, X_n\}$ of non-zero prices is normalised by dividing the complete time series by its first number to yield $\{1, \frac{X_2}{X_1}, \frac{X_3}{X_1}, \dots, \frac{X_n}{X_1}\}$. The same normalisation has been done before calculating the distances with other measures also.

Now, suppose $\{x_t\}_{t \in [1, \dots, n]}$ and $\{y_t\}_{t \in [1, \dots, n]}$ be the normalised prices of two stocks. Then, the SSD [7] is given by:

$$SSD_{X,Y} = \frac{1}{n} \sum_{t=1}^n (x_t - y_t)^2$$

In the highly cited work by Gatev et al [7], it has been observed that the pairs which have lower SSD distance between them are better candidates for pairs trading. They also show that SSD metric is the best for this task. Still, it may not always be the case. Chen et al [19] empirically and theoretically argued that correlation selection metric can show better performance than SSD measure.

The following is a brief description of the CCT measure, which has also been used in experimentation.

3. Cross-Correlation Type Measure

Given two-time series X_T and Y_T (each of length T), here the interest is in finding a dissimilarity measure between them. Let ' m ' be the maximum value of lead or lag being taken into consideration for the calculation of dissimilarity between the two-time series. We consider a segment of time series X_T starting from ' $m+1$ ' and ending at ' $T-m-1$ ', and divide this segment into ' n ' equal parts each of length ' p '. Here, we conveniently choose ' m ', ' n ' and p such that $2m+np+1 = T$. This is required to make sure that all the data points in time series are utilized for the calculation of dissimilarity measure. Though it would still suffice if ' $2m+np+1$ ' is slightly less than T . Another couple of restrictions are that $m \leq p$ and $p \geq 15$. These restrictions have been imposed to avoid unwanted cross-correlations. Now, with this background we define our CCT similarity measure, which is given by the following:

$$CCT = \frac{1}{n} \times \sum_{l=1}^n \max\{CCT_k(m+1+p \times (l-1)) \mid -m \leq k \leq m\} \quad (1)$$

where $CCT_k(i)$ is defined as follows

$$CCT_k(i) = \frac{\sum_{j=i}^{i+p} (x_{j+1} - x_j)(y_{j+k+1} - y_{j+k})}{\sqrt{\sum_{j=i}^{i+p} (x_{j+1} - x_j)^2} \sqrt{\sum_{j=i}^{i+p} (y_{j+k+1} - y_{j+k})^2}}$$

The value of CCT_k is the same as the correlation between returns of a segment of time series Y_t at a lead k with respect to a segment of time series X_t . Further information about this metric can be found in [8].

3. Proposed DCCT Measure

The proposed distance measure is described in this section. Let x_t, y_t ($t \in 1, 2, \dots, n$) be two time-series of normalised prices of two stocks. Then the DCCT's computation requires the use of an alignment path.

Alignment path is a sequence $P = (P_1, P_2, \dots, P_l, \dots, P_L)$ with $P_l = (p_l, q_l) \in [1:n]^2$ for $l \in [1:L]$ satisfying the following conditions:

- (i) Boundary Condition: Given a parameter ψ , (say 100), then $p_1 \leq \psi$, $q_1 \leq \psi$, $p_L \geq n - \psi$ and $q_L \geq n - \psi$.
- (ii) Monotonicity Condition: $p_1 \leq p_2 \leq \dots \leq p_L$ and $q_1 \leq q_2 \leq \dots \leq q_L$.
- (iii) Step size Condition: $P_{l+1} - P_l \in \{(1,0), (0,1), (1,1)\}$ for $l \in [1:L-1]$.

The parameter name ' ψ ' (which stands for 'Post Suffix Invariant'), has been inspired from [14] where they introduced this parameter to relax the boundary condition in DTW.

The alignment path is computed using 'Dynamic programming' techniques as done in DTW measure [14].

Now, the DCCT measure can be defined using the above definition. Let us denote the computation of the alignment path by step 1.

Step 1: Computation of Alignment path

Let $CR(p_i, q_i)$, a function over the sequence $P_l = (p_l, q_l)$, be defined as follows:

$$CR(p_i, q_i) = \frac{\sum_{j=i}^{i+p} (x_{p_j} - x_{p_{j+1}})(y_{q_j} - y_{q_{j+1}})}{\sqrt{\sum_{j=i}^{i+p} (x_{p_j} - x_{p_{j+1}})^2} \sqrt{\sum_{j=i}^{i+p} (y_{q_j} - y_{q_{j+1}})^2}},$$

Here, ‘p’ is the window size parameter. It is analogous to the parameter ‘p’ defined in equation (1) corresponding to CCT measure. In our empirical analysis, we pick the same values of ‘p’ for both CCT and DCCT measure in each experiment.

Then, we find an alignment path $P = (P_1, P_2, \dots, P_L)$ which maximises the function $\sum_{i=1}^L CR(p_i, q_i)$ given the constraints as mentioned earlier. This is done through Dynamic Programming techniques as used in DTW measure.

Step 2: Computation of DCCT measure

Finally, DCCT measure is the summation of function CR over the alignment path, calculated in the last step. Formally,

$$DCCT = \max_{(p_i, q_i), i=1 \dots L} \sum_{i=1}^L CR(p_i, q_i)$$

where, maximisation is done over an alignment path (p_i, q_i) .

In the present experiments, the parameter ‘psi’ has been kept equal to the parameter ‘p’ just for simplicity. Notice, DCCT measure has a striking resemblance with CCT distance measure. Figure 1 explains the intuition behind the DCCT measure.

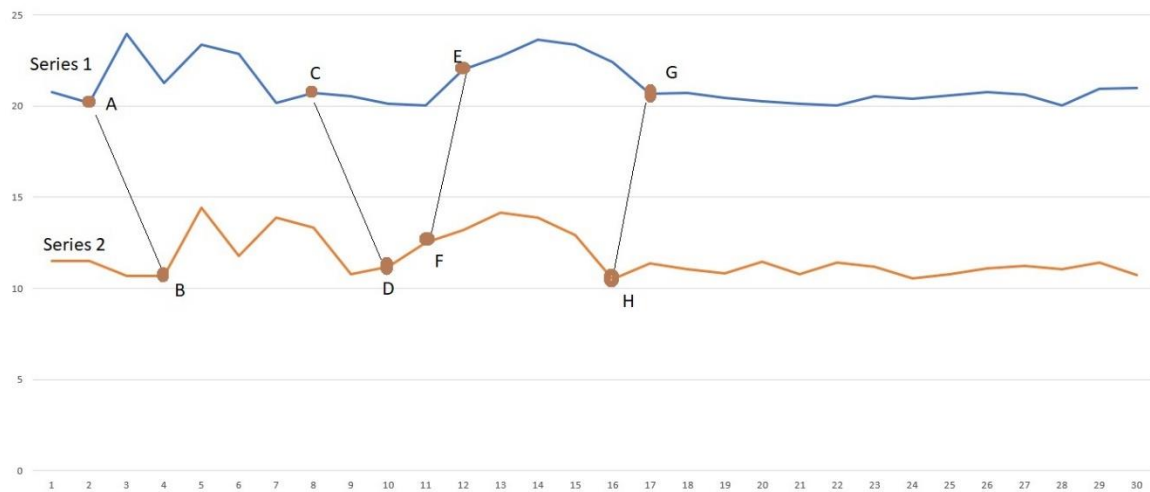


Figure 1: Series 1 lags from series 2 by a value of 2 in the region A to C. In the region E to G, series 1 leads series 2 by a value of 1. While calculating the distance between the two series, CCT measure will allow a change in lead-lag value in different parts of the time series, but the changes may be abrupt or huge. Whereas, DCCT measure will allow the lead-lag value to change continuously from -2 to 1 while calculating its final measure value. So, it will be able to incorporate such subtle changes in its final value.

Identifying the pairs by DCCT

- 2) Convert closing price to normalised price series
- 3) For every two normalised price series of randomly picked two stocks.
 - i. find the optimal path for lead-lag values by Dynamic programming technique
 - ii. over the optimal path, calculate the DCCT measure
- 4) pick the pair with highest DDCT measure

Figure 2: Brief steps of DCCT measure.

4. Methodology and Data description

4.1 Methodology

Once, the stock pairs have been identified for pairs trading, a pair trading strategy was used for trading in the market. The pair trading strategy was developed using python package backtrader [20]. Specifically, the package contains an example pairs trading script, which was conveniently modified to suit the experiments. The code for the experiments could be found at the website ‘www.github.com/kartikay94’. The dataset may also be requested from the corresponding author.

Three different trigger points were used for entering into positions. These trigger points are calculated as follows.

- 1) First, the spread is calculated using moving regression. More specifically spread at point t is given by $\epsilon_t = Y_t - s \times X_t - c$. Here slope ‘ s ’ and intercept ‘ c ’ parameters are determined using regression on past ten values at each time stamp.
- 2) Then, using these spread values, z-score is calculated with a moving window of period P .
- 3) Now, we describe the calculation of z- score. Suppose it is required to find z-score at T^{th} time stamp value of 2 price series $\{X_t\}_{t \in [1, \dots, n]}$ and $\{Y_t\}_{t \in [1, \dots, n]}$. Then,

$$z - score = \frac{\epsilon_T - \frac{\sum_{t=T-P}^{T-1} \epsilon_t}{P}}{\sqrt{\frac{(\sum_{t=T-P}^{T-1} (\epsilon_t - \frac{\sum_{t=T-P}^{T-1} \epsilon_t}{P}))^2}{P}}}$$

where $\epsilon_i = Y_i - s \times X_i - c$, and P is some pre-determined period.

- 4) Z-score is calculated corresponding to three different periods i.e., 14, 21 and 35.

These different moving-window periods correspond to the three different trigger points which are experimented upon.

- 5) Here it takes appropriate positions when the absolute value of z-score rises up to the value two. Specifically, it buys the underperforming stock and sells the better-performing stock. This is done in anticipation that stocks will again fall back to equilibrium relationship. These positions are neutralised when z-score crosses zero. The positions are also neutralized at the end of the trading session.

The pairs trading strategy experimented upon is a standard strategy. The use of z-score for generating trading signals can be seen in [21]. Also, the z-score's threshold value of two for entering into position is earlier employed in [7]. Figure 3 gives an outline of this pairs trading strategy.

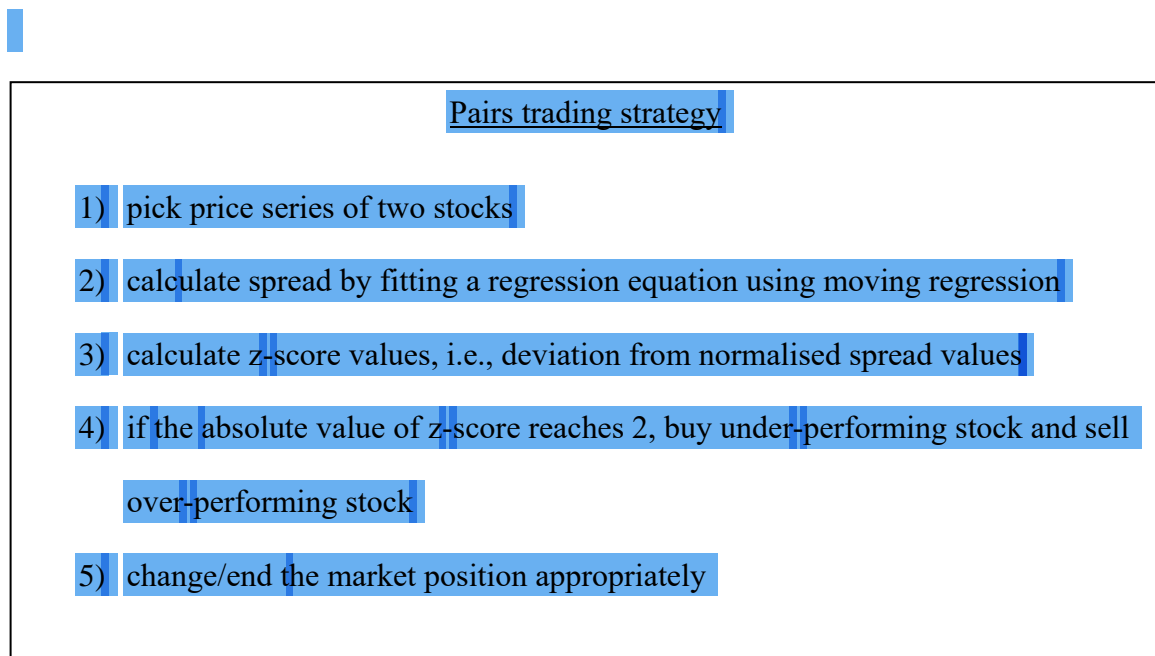


Figure 3: Brief overview of steps of pairs trading strategy.

Performance of different pairs is judged through ‘Profit margin’ as shown in Table 2,3. Though such pair trading strategies should be self-financing, but for the sake of evaluation, ‘profit margin’ is considered which is defined as follows. Suppose, we start with a capital of 1,00,000. We enter into the market by buying 50,000 worth of shares and short-selling shares of the equal amount. The broker keeps the whole amount 50,000 obtained through short-selling as margin money. Once the positions need to be closed, we obtain a principal amount by selling off the bought shares and broker’s release of the left-over margin money. If a new position needs to be entered after that, this capital is divided into two. Then that much amount of shares are bought and sold simultaneously. This technique is repeated till the end of the trading window when all positions are closed and final capital is calculated. Finally, we calculate the profit margin as the percentage of final profit over the initial worth of shares bought (50,000). This evaluation metric is also used in [22]. Figure 4 briefly presents the steps involved in the overall experiments.

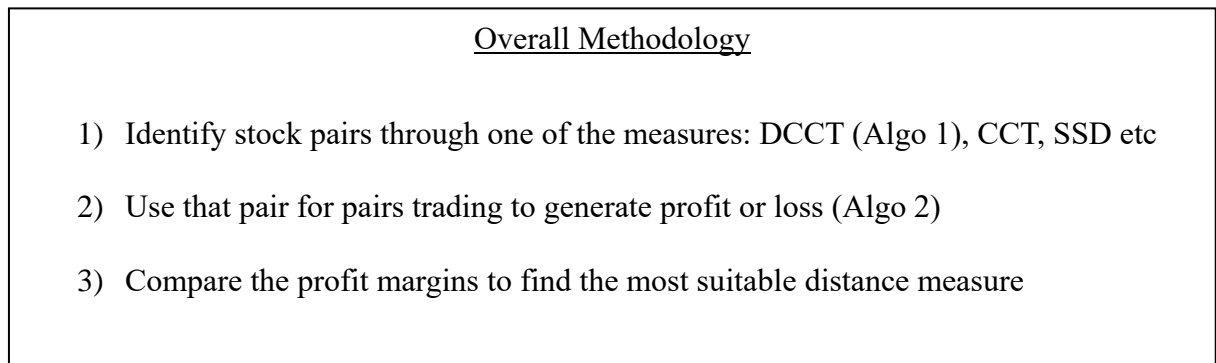


Figure 4: Outline of the overall methodology.

4.2 Data Description

The experiments are conducted on two datasets. The first dataset is taken from Dow Jones Industrial Average index and its constituents. The timespan of the data is 2000

trading days starting from 2008-08-25 to 2016-08-03. The second dataset comprised of Indian stocks. It comprised of 2000 trading days prices beginning from 5 Aug 2008 till 13 April 2016. Table 1 gives the name of the 28 Indian companies whose stock prices were used in experimentation. The datasets may also be made available on request from the corresponding author.

Indian Companies forming Dataset 2					
1	SBI Bank	11	HDFC	20	INDIAN OIL
2	Baroda	12	ICICI	21	ONGC
3	PNB	13	AXIS	22	BHARAT
4	IDBI Ltd	14	KOTAK	23	ESSAR
5	Central Bank I	15	INDUSIND	24	Cairn I L
6	Canara Bank	16	Yes Bank Ltd	25	Hindustan P C
7	UBI	17	Federal Bank Ltd	26	Aban Offshore
8	Bank Of India	18	Karur Vysya	27	Hind Oil exp
9	Syndicate	19	South Indian B L	28	Manga Refinery
10	Indian Bank				

Table 1: Acronyms of the Indian companies belonging to Dataset 2.

Taking inspiration from [18] we adopted the rolling windows approach with 500 trading days for identifying the best pairs for trade, and the next 250 trading days are for conducting trading on the selected pairs. So, overall we obtain 6 windows for training purpose and 6 windows for testing purpose. Thus, the timespan is divided into 6 Cross Validation (CV) windows and profit is evaluated over these different periods.

Two metrics, namely Pearson correlation and SSD, are used for comparison with the CCT measure and proposed DCCT measure.

1. Empirical Analysis

The experiments are done with three parameter settings for both the metrics DCCT and CCT. DCCT has the following parameters which need to be predetermined: ‘p’ and ‘psi’.

Here ‘psi’ refers to the deviation from the endpoints allowed in the alignment path. ‘p’ refers to the window size or size of minor time series, which is used for calculating changing lead-lag relationships. For simplicity, the two parameter values are kept the same in the present paper. The three parameter settings used for DCCT are {(psi=100), (psi=50), (psi=25)}. CCT has the following parameters which need to be pre-determined ‘p’, ‘m’ and ‘n’. The parameter settings used for CCT are {(p=100) , (p=50), (p=25) }. Here, the maximum lead-lag value (m) is made equal to ‘p’. The parameter ‘n’ is determined according to the values of other parameters and the constraints in the parameter settings as described in section 2.

The tables below give the result of the experiments. The leftmost column ‘No. of pairs’ gives the number of top pairs used for trading separately to generate a combined profit. This profit is evaluated as ‘profit margin’ (described earlier) and is given under the column ‘Value’ for each metric. The column ‘Rank’ gives the rank of the ‘profit margin’ as compared with other measures’ performance within each row.

Table 2: Results corresponding to DJIA dataset with different values of period (P).

a) Period (P) = 14

No. of pairs	DCCT						CCT						CORR		SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25		CORR		SSD	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	-0.463	6	-0.778	7	0.027	1	-0.450	5	-0.207	3	-0.252	4	-0.133	2	-0.877	8
4	-0.539	6	-0.689	7	0.047	2	-0.140	3	-0.387	5	-0.326	4	0.220	1	-0.845	8
6	-0.381	5	-0.445	6	-0.149	4	-0.099	3	-0.501	7	-0.083	2	0.075	1	-0.520	8
8	-0.150	3	-0.424	8	-0.190	4	-0.190	5	-0.285	6	-0.103	2	0.005	1	-0.391	7
10	-0.368	6	-0.717	8	0.041	2	-0.306	5	-0.403	7	-0.225	4	-0.048	3	0.225	1
15	-0.070	3	-0.285	6	-0.080	4	-0.523	8	-0.130	5	-0.353	7	0.036	2	0.935	1
20	-0.008	4	-0.245	7	-0.011	5	-0.380	8	0.154	2	-0.192	6	0.011	3	0.676	1
30	-0.044	5	-0.320	7	0.008	3	-0.405	8	0.024	2	-0.042	4	-0.067	6	0.499	1

b) Period (P) = 21

No. of pairs	DCCT						CCT						CORR		SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25		CORR		SSD	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	-0.202	4	-0.377	6	0.121	1	-0.485	7	-0.255	5	0.049	2	0.014	3	-0.778	8
4	-0.213	3	-0.470	6	0.000	1	-0.359	5	-0.331	4	-1.672	8	-0.091	2	-1.123	7
6	0.307	1	-0.132	4	-0.121	3	0.216	2	-1.343	7	-0.976	6	-1.377	8	-0.712	5
8	-0.285	2	-0.601	4	-0.101	1	-0.650	5	-1.202	8	-0.676	6	-0.967	7	-0.519	3
10	-0.471	3	-0.855	7	0.110	1	-0.594	4	-1.202	8	-0.668	5	-0.790	6	-0.369	2
15	-0.197	2	-0.713	7	-0.018	1	-0.639	6	-0.816	8	-0.556	5	-0.430	4	-0.390	3
20	-0.139	2	-0.622	7	-0.031	1	-0.682	8	-0.490	5	-0.338	3	-0.419	4	-0.578	6
30	-0.040	2	-0.428	5	0.037	1	-0.784	8	-0.360	4	-0.429	6	-0.353	3	-0.623	7

c) Period(P) = 35

No. of pairs	DCCT						CCT						CORR		SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25		CORR		SSD	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	-0.101	5	0.075	1	-0.045	4	-0.363	8	-0.329	7	0.039	2	0.013	3	-0.242	6
4	-0.259	6	-0.154	3	0.014	1	-0.069	2	-0.239	5	-1.110	8	-0.184	4	-0.537	7
6	0.304	2	-0.248	3	-0.418	5	0.521	1	-1.273	7	-0.424	6	-1.323	8	-0.366	4
8	-0.359	3	-0.635	6	-0.390	4	-0.303	2	-1.311	8	-0.287	1	-0.950	7	-0.575	5
10	-0.352	4	-0.712	6	-0.040	1	-0.319	3	-1.251	8	-0.203	2	-0.789	7	-0.398	5
15	-0.023	1	-0.518	6	-0.113	2	-0.602	7	-0.806	8	-0.299	3	-0.488	5	-0.456	4
20	-0.050	1	-0.505	6	-0.116	2	-0.686	7	-0.906	8	-0.122	3	-0.404	5	-0.347	4
30	-0.185	2	-0.364	6	-0.079	1	-0.707	8	-0.605	7	-0.290	4	-0.338	5	-0.214	3

From Table 2, we observe SSD and CORR metrics performance has been not good especially in results corresponding to period 21 and 35. Also, DCCT (psi = 25) has given consistently better results than these two measures. This prompted us to find a measure which combines the power of SSD with the proposed measures. This is because the DCCT measure is intrinsically related to the correlation measure. SSD being totally unrelated measure should be able to add significant information to the DCCT measure.

The following experiments are conducted to showcase the effectiveness of combined measures as compared to just SSD. Here the DCCT, the CCT and the correlation measure, are combined with the SSD measure. Then, the dual objectives are optimised to find the best pairs. The pairs are sorted according to the non-dominated front while

maximising the dual objectives. The pairs lying in each front are finally sorted according to SSD. Thus, a sorted list of pairs is obtained.

Table 3: Results corresponding to DJIA dataset with different values of period (P).

a) Period (P) = 14

No. of pairs	DCCT & SSD						CCT & SSD						CORR & SSD		SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25		CORR & SSD		SSD	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	-0.736	7	-0.736	5	-0.736	6	-0.613	3	-0.616	4	-0.345	2	0.506	1	-0.877	8
4	-0.248	4	-0.253	5	-0.258	6	-0.216	3	-0.102	2	-0.024	1	-0.463	7	-0.845	8
6	-0.040	7	-0.025	6	-0.014	5	0.063	3	0.076	1	0.066	2	0.057	4	-0.520	8
8	-0.146	6	-0.212	7	-0.064	3	-0.020	2	-0.133	5	-0.079	4	0.106	1	-0.391	8
10	-0.105	7	-0.056	6	-0.033	3	-0.041	4	-0.150	8	-0.053	5	0.056	2	0.225	1
15	0.012	7	0.065	5	0.094	2	0.043	6	0.005	8	0.073	4	0.082	3	0.935	1
20	0.129	4	0.074	6	0.160	2	0.087	5	-0.022	8	0.138	3	0.018	7	0.676	1
30	0.021	8	0.108	5	0.098	6	0.276	2	0.192	3	0.188	4	0.023	7	0.499	1

b) Period (P) = 21

No. of pairs	DCCT & SSD						CCT & SSD						CORR & SSD		SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25		CORR & SSD		SSD	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	-0.422	7	-0.422	5	-0.422	6	-0.408	4	-0.282	3	-0.126	2	0.329	1	-0.778	8
4	-0.354	5	-0.336	3	-0.354	4	-0.364	6	-0.183	1	-0.194	2	-0.428	7	-1.123	8
6	0.289	1	0.029	5	-0.036	6	0.090	4	0.102	3	0.104	2	-0.051	7	-0.712	8
8	0.324	1	-0.648	6	-0.719	7	-0.597	3	-0.628	5	-0.605	4	-0.728	8	-0.519	2
10	-0.233	1	-0.431	3	-0.509	6	-0.513	7	-0.462	4	-0.503	5	-0.528	8	-0.369	2
15	-0.374	2	-0.423	4	-0.361	1	-0.563	8	-0.560	7	-0.546	6	-0.518	5	-0.390	3
20	-0.241	2	-0.232	1	-0.303	5	-0.257	3	-0.349	6	-0.279	4	-0.451	7	-0.578	8
30	-0.169	3	-0.150	2	-0.140	1	-0.315	7	-0.170	4	-0.190	5	-0.233	6	-0.623	8

c) Period (P) = 35

No. of pairs	DCCT & SSD						CCT & SSD						CORR & SSD		SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25		CORR & SSD		SSD	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	-0.135	6	-0.135	4	-0.135	5	-0.142	7	-0.024	3	0.201	2	0.303	1	-0.242	8
4	-0.249	6	-0.168	3	-0.187	5	-0.300	7	-0.086	1	-0.104	2	-0.182	4	-0.537	8
6	0.427	1	0.123	5	0.162	2	0.118	6	0.124	4	0.132	3	0.077	7	-0.366	8
8	0.273	1	-0.084	7	0.087	2	-0.018	4	-0.019	5	-0.075	6	0.073	3	-0.575	8
10	0.221	1	-0.039	4	-0.115	7	-0.032	2	-0.034	3	-0.092	5	-0.093	6	-0.398	8
15	-0.055	1	-0.114	2	-0.144	3	-0.364	7	-0.215	4	-0.255	6	-0.236	5	-0.456	8
20	-0.392	5	-0.196	2	-0.330	3	-0.522	7	-0.399	6	-0.096	1	-0.524	8	-0.347	4
30	-0.374	5	-0.424	6	-0.463	7	-0.199	1	-0.280	4	-0.254	3	-0.513	8	-0.214	2

Table 3 gives the results corresponding to DJIA dataset for three types of triggers corresponding to periods 14, 21 and 35. In table 3 (b) and 3 (c), it can be seen that the SSD and CORR&SSD consistently show the worst performance as compared to the other metrics. Their performance in table 3 (a) is also poor. In contrast ‘DCCT (psi=100) & SSD’ has given very good results specifically corresponding to periods 21 and 35.

Also ‘DCCT(psi=25) & SSD’ and ‘CCT (p=25) & SSD’ has shown better performance than the two rightmost metrics.

Another important factor which is intuitive is also observed in these tables. The performance of SSD measure is mostly the worst. This is because other metrics already combine the power of SSD measure, in their measure value. Thus, in the next set of experiments, SSD performance is not evaluated.

Indian Data

Then experiments are also done corresponding to the second dataset of Indian companies. Table 4 give the results. In Table 4 (b) and (c), we observe that CORR & SSD show mostly the worst performance amongst all other metrics. Its performance in table 4 (a) is also poor. In contrast the performance of the measure ‘DCCT & SSD(psi=50) is far better than the rightmost ‘Corr & SSD’ metric. ‘CCT(p=100) & SSD’ has also given consistently better results than ‘Corr & SSD’.

Table 4: Results corresponding to Indian dataset with different values of period (P).

a) Period (P) = 14

No. of pairs	DCCT & SSD						CCT & SSD						CORR & SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25			
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	0.225	8	0.796	5	1.922	1	0.687	6	1.196	4	0.500	7	1.196	3
4	-0.319	8	0.477	2	-0.206	7	0.379	4	0.357	5	0.275	6	0.442	3
6	0.442	5	0.466	4	0.254	7	0.548	2	0.521	3	0.156	8	0.388	6
8	0.285	4	0.394	1	0.200	6	0.239	5	0.104	7	0.373	2	-0.031	8
10	0.191	5	0.309	2	-0.199	8	0.240	3	0.133	6	0.397	1	-0.127	7
15	0.134	2	-0.037	7	-0.108	8	0.026	5	0.076	3	0.186	1	0.045	4
20	-0.091	5	-0.178	8	0.129	1	-0.040	3	-0.042	4	0.008	2	-0.101	6
30	-0.250	8	-0.228	7	0.154	1	-0.201	5	-0.098	3	-0.161	4	-0.075	2

b) Period (P) = 21

No. of pairs	DCCT & SSD						CCT & SSD						CORR & SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25			
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	0.180	8	1.012	5	1.317	2	0.938	6	1.052	4	0.735	7	1.052	3
4	-0.537	7	0.157	4	-0.545	8	0.177	3	0.059	5	-0.204	6	0.284	2
6	-0.086	7	0.182	3	0.051	6	0.293	2	0.080	5	-0.274	8	0.116	4
8	-0.138	6	0.280	1	0.035	4	0.100	2	-0.171	7	0.042	3	-0.250	8
10	-0.118	4	0.071	1	-0.309	7	0.000	2	-0.150	5	-0.094	3	-0.353	8
15	-0.147	2	-0.292	4	-0.362	7	-0.144	1	-0.254	3	-0.333	5	-0.420	8
20	-0.231	3	-0.358	7	-0.104	1	-0.212	2	-0.306	4	-0.321	6	-0.421	8
30	-0.425	8	-0.239	3	-0.067	1	-0.293	7	-0.285	6	-0.247	5	-0.241	4

c) Period (P) = 35

No. of pairs	DCCT & SSD						CCT & SSD						CORR & SSD	
	psi = 100		psi = 50		psi = 25		p=100		p=50		p=25			
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
2	0.609	5	0.651	4	0.775	2	0.713	3	0.429	8	0.478	6	0.429	7
4	0.123	5	0.257	4	-0.394	8	0.409	2	0.091	6	-0.019	7	0.265	3
6	0.363	3	0.209	5	0.250	4	0.521	2	0.089	7	-0.108	8	0.109	6
8	0.365	1	0.341	2	0.188	5	0.247	4	0.054	6	0.272	3	-0.007	8
10	0.176	2	0.263	1	-0.104	7	0.074	5	0.112	4	0.123	3	-0.148	8
15	0.157	1	-0.124	6	-0.120	5	0.031	2	0.015	3	-0.129	7	-0.108	4
20	-0.054	3	-0.185	7	-0.010	2	-0.071	4	-0.077	5	-0.085	6	-0.245	8
30	-0.227	8	-0.023	3	0.006	2	-0.158	7	-0.088	6	-0.079	4	-0.088	5

2. Discussion

The paper showed that for a standard strategy, the proposed metrics with different parameters all showed better or comparable results than existing metrics. But still,

certain important things, pertaining to parameter settings, faster computation techniques and possible generalisations are discussed in this section.

The DCCT and CCT measure may be efficiently computed using the convolution theorem [23] or even more efficiently through Fast Fourier Transform algorithm [24], [25].

DCCT measure significantly generalises DTW measure, and this form of generalisation can allow further modifications. DTW works with Euclidean distance which is from point to point distance. But measures like correlation etc can only be used to find distances between two 'sets of points'. When dealing with such distance measures, this proposed methodology of DCCT measure may be utilised by slight modifications. That is, the CR measure here may be changed to any other measure to obtain a new desired measure.

Here we discuss optimal parameter-setting for parameter 'p' in DCCT. The parameter 'p' should be smaller than the complete length of time series. As otherwise, the DCCT measure will reduce to simple correlation measure. Secondly, it should not be too small. This is because two smaller time series have a higher probability of having a spurious high correlation between them. This observation can be justified theoretically as well. As reported in [26], more the sample size (or length of the time series) more accurate is the empirically determined non-zero coefficient value. This means longer the length, lesser are the chances of a spuriously high correlation between the two-time series. In other words, if the hypothesis is that the two-time series are uncorrelated, then this hypothesis is more strongly rejected when the number of samples is more, for the same empirically-determined non-zero correlation value. Thus, the probability of the existence of spurious correlation decreases as the length of time series increases. This indicates that the

parameter 'p' should be significantly smaller than the whole time series but should not be too small as then the correlation value becomes spurious.

Further experiments may be done to formulate ways of finding the best parameter, which may be through cross-validation, or analysis of past time series, or some an average of measures with different parameters. Also, a combination of these metrics can be chosen to identify the final pairs. This combination can be done in lots of ways, where dual or higher objectives are optimised. Also, other strategies could be devised which are suitable to take advantage of any lead-lag relationship between the two stocks. As an instance, suppose we know stock A is correlated at a lead of 1 from stock B. So, we may take positions in futures derivatives of stock A and spot market in stock B. Thus, there exist an opportunity to extract profit through the lead-lag information further.

This study primarily concerns with the selection metric and not about pair trading strategies. Thus an analysis of each run of the strategy in terms of profitable positions, loss-making positions, trigger points etc, which warrants further attention, has been deferred for a later study. Also, the experiments here pertain to strategies trading on a daily basis. Similar experiments could be conducted for higher frequency trading strategies like HFT. These shorter interval time series (tick by tick or minute-wise data) are more susceptible to measures like correlation and the corresponding 'lead-lag effect'.

In future, experiments could be done to ascertain the effect of window parameter 'p' on the lead-lag relationship in DCCT metric. One may be interested in understanding the effect of change in lead-lag value at any one point, as the parameter 'p' changes. That is, the empirically determined temporal delay may shorten or lengthen based on the value of

parameter 'p' of DCCT. Also, what such changes indicate about the stock pairs. Such analysis, which requires extensive computation, has been deferred for a future study.

The best metric may further depend on strategy and data type. For a pairs trading strategy, the measures like 'CCT' and 'DCCT' should also be looked upon before choosing relevant pairs.

Finally, these experiments do show the effectiveness of the proposed measures over the typically used measures. Pair selection is a very subjective matter and requires extensive evaluation of businesses, management, intended mergers and acquisitions etc. But overall, these metrics will be an aid for the financial traders.

3. Conclusion

The present study proposed one distance measure for selecting pairs for pairs trading. The distance measure can effectively incorporate the changing lead-lag relationship between the stocks. The different parameter settings for the proposed measure are compared with other measures. Experiments are done on two different data sets. The results indicate the usefulness of the proposed measure. Stock pairs identified through optimising dual measures (SSD and the proposed DCCT measure) consistently generate higher profit. The work concludes that the lead-lag relationship between the stocks is important and can be used for identifying pairs for trading.

References

- [1] C. Krauss, “STATISTICAL ARBITRAGE PAIRS TRADING STRATEGIES: REVIEW AND OUTLOOK,” *J. Econ. Surv.*, vol. 31, no. 2, pp. 513–545, 2017.
- [2] S. Mallaby, “Learning to love hedge funds,” *Wall Street Journal*, 2010.
- [3] M. Mitchell and T. Pulvino, “Characteristics of Risk and Return in Risk Arbitrage,” *J. Finance*, vol. 56, no. 6, pp. 2135–2175, Dec. 2001.
- [4] R. A. Korajczyk and R. Sadka, “Are Momentum Profits Robust to Trading Costs?,” *J. Finance*, vol. 59, no. 3, pp. 1039–1082, Jun. 2004.
- [5] B. D. Grundy and J. S. M. Martin, “Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing,” *Rev. Financ. Stud.*, vol. 14, no. 1, pp. 29–78, Jan. 2001.
- [6] D. Wen, C. Ma, G. J. Wang, and S. Wang, “Investigating the features of pairs trading strategy: A network perspective on the Chinese stock market,” *Phys. A Stat. Mech. its Appl.*, 2018.
- [7] E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst, “Pairs Trading: Performance of a Relative-Value Arbitrage Rule,” *Rev. Financ. Stud.*, vol. 19, no. 3, pp. 797–827, Oct. 2006.
- [8] K. Gupta and N. Chatterjee, *Financial time series clustering*, vol. 84. 2018.
- [9] M. Hoffmann, M. Rosenbaum, and N. Yoshida, “Estimation of the lead-lag parameter from non-synchronous data,” *Bernoulli*, vol. 19, no. 2, pp. 426–461, Mar. 2013.

- [10] F. de Jong and T. Nijman, “High frequency analysis of lead-lag relationships between financial markets,” *J. Empir. Financ.*, vol. 4, no. 2–3, pp. 259–277, Jun. 1997.
- [11] J. Kang, C. J. Lee, and S. Lee, “An Empirical Investigation of the Lead-Lag Relations of Returns and Volatilities among the KOSPI200 Spot, Futures and Options Markets and their Explanations,” *J. Emerg. Mark. Financ.*, vol. 5, no. 3, pp. 235–261, Dec. 2006.
- [12] C. Y. Robert and M. Rosenbaum, “On the limiting spectral distribution of the covariance matrices of time-lagged processes,” *J. Multivar. Anal.*, vol. 101, no. 10, pp. 2434–2451, Nov. 2010.
- [13] D. J. Berndt and J. Clifford, “Using dynamic time warping to find patterns in time series,” *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*. AAAI Press, pp. 359–370, 1994.
- [14] D. F. Silva, G. E. A. P. A. Batista, and E. Keogh, “On the Effect of Endpoints on Dynamic Time Warping,” 2016.
- [15] T. Felix, “Dynamic Programming Algorithms in Speech Recognition,” *Rev. Inform. Econ.*, vol. 2(46), 2008.
- [16] Q. Zhu, G. Batista, T. Rakthanmanon, and E. Keogh, “A Novel Approximation to Dynamic Time Warping allows Anytime Clustering of Massive Time Series Datasets,” in *Proceedings of the 2012 SIAM International Conference on Data Mining*, Philadelphia, PA: Society for Industrial and Applied Mathematics, 2012, pp. 999–1010.

- [17] Jieren Wang, C. Rostoker, and A. Wagner, “A high performance pair trading application,” in *2009 IEEE International Symposium on Parallel & Distributed Processing*, 2009, pp. 1–8.
- [18] J. P. Ramos-Requena, J. E. Trinidad-Segovia, and M. A. Sánchez-Granero, “Introducing Hurst exponent in pair trading,” *Phys. A Stat. Mech. its Appl.*, vol. 488, pp. 39–45, Dec. 2017.
- [19] H. (Jason) Chen, S. J. Chen, F. Li, and F. Li, “Empirical Investigation of an Equity Pairs Trading Strategy,” *SSRN Electron. J.*, Apr. 2012.
- [20] “backtrader 1.9.69.122 documentation.” [Online]. Available: <https://www.backtrader.com/docu/index.html>. [Accessed: 11-Apr-2019].
- [21] K. Belkhat, “Statistical Arbitrage: Systematic Equity Pairs Trading,” *ETH Zurich*, 2017. [Online]. Available: https://www.researchgate.net/publication/317598743_Statistical_Arbitrage_Systematic_Equity_Pairs_Trading. [Accessed: 29-May-2019].
- [22] B. Quinn, A. Hanna, and F. MacDonald, “Picking up the pennies in front of the bulldozer: The profitability of gilt based trading strategies,” *Financ. Res. Lett.*, vol. 27, pp. 214–222, Dec. 2018.
- [23] Y. Katznelson, *An introduction to harmonic analysis*. Cambridge University Press, 2004.
- [24] J. W. Cooley and J. W. Tukey, “An algorithm for the machine calculation of complex Fourier series,” *Math. Comput.*, vol. 19, no. 90, pp. 297–297, May 1965.

- [25] J. Paparrizos and L. Gravano, “Fast and Accurate Time-Series Clustering,” *ACM Trans. Database Syst.* *ACM Trans. Database Syst*, vol. 42, no. 49, 2017.
- [26] M. A. Bujang and N. Baharum, “Sample Size Guideline for Correlation Analysis,” *World J. Soc. Sci. Res.*, vol. 3, no. 1, p. 37, Mar. 2016.