Journal of the Operations Research Society of Japan
Vol. 60, No. 3, July 2017, pp. 244–261

OPTIMAL MULTIPLE PAIRS TRADING STRATEGY USING DERIVATIVE FREE OPTIMIZATION UNDER ACTUAL INVESTMENT MANAGEMENT CONDITIONS

Rei Yamamoto
Musashi University

Norio Hibiki
Keio University

(Received June 7, 2016; Revised August 17, 2016)

Abstract Pairs trading strategy has a history of at least 30 years in the stock market and is one of the most common trading strategies used today due to its understandability. Recently, Yamamoto and Hibiki [13] studied optimal pairs trading strategy using a new approach under actual fund management conditions, such as transaction costs, discrete rebalance intervals, finite investment horizons and so on. However, this approach cannot solve the problem of multiple pairs because this problem is formulated as a large scale simulation based non-continuous optimization problem. In this research, we formulate a model to solve an optimal pairs trading strategy problem using multiple pairs under actual fund management conditions. Furthermore, we propose a heuristic algorithm based on a derivative free optimization (DFO) method for solving this problem efficiently.

Keywords: Finance, pairs trading strategy, multiple pairs, fund management conditions, derivative free optimization

1. Introduction

Pairs trading strategy has been applied for at least 30 years in the stock market and currently among the most commonly used trading strategies because of its understandability (Vidyamurthy [11]).

Known as a statistical arbitrage, pairs trading strategy analyzes the difference between two stock prices. The investor selects two correlated stocks and monitors the difference (or spread) between their stock prices. When the spread is wider than a certain threshold, the investor buys the undervalued(i.e. the lower-priced) stock and sells the overvalued(i.e. the higher-priced) stock, because the spread should eventually converge to a normal level. After convergence, the trade is closed out by taking opposite positions in the stocks.

As an example, Figure 1 shows the spread of Toyota Motor and Toyota Industries in 2012. Here, the spread is the difference between two logarithmic prices. The solid line indicates the spread, while the dotted and broken lines show the strategy's opening and closing thresholds, specified as 1.0 and 0.3 standard deviations from the mean, respectively.

According to this figure, the spread is mean-reverting. The investor bought the stock of Toyota Industries and sold the stock of Toyota Motor when the spread reached 0.39 (exceeding the opening threshold) on May 16. The investor unwound this position when the spread converged to 0.36 (within the closing threshold) on June 11 and, thus, received a 3%(0.39 − 0.36) return.

The profit obtained from the pairs trading strategy derives from the cointegration of two stock prices. Cointegration is a statistical property. Stocks are said to be cointegrated when their individual price processes are nonstationary, while the price process of their com-

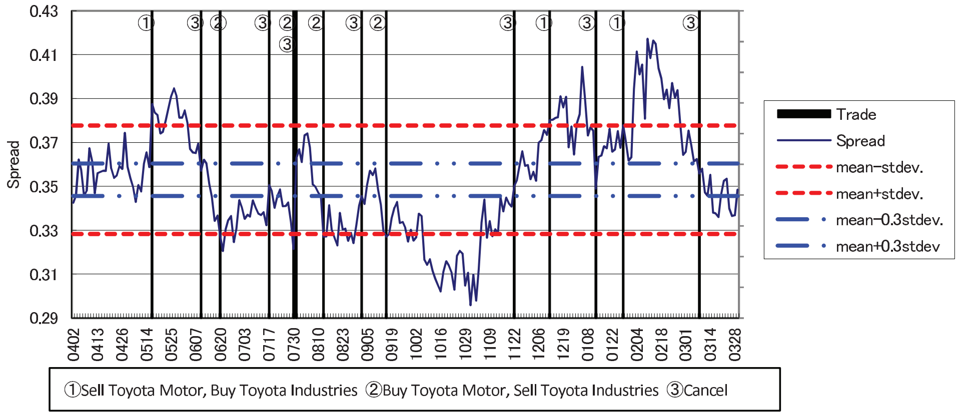

Figure 1: Spread of Toyota Motor and Toyota Industries in 2012

mon portfolio is stationary. The pairs trading strategy is successful because cointegration incorporates mean reversion into the spread.

The theoretical and empirical aspects of pairs trading have been extensively studied. In a theoretical evaluation, Elliott et al. [3] indicated that the above-discussed trading scheme, with reference to the spread and threshold positions, optimizes the pair trading strategy. They also optimized the threshold by considering the statistical property of cointegrated pairs.

Mudchanatongsuk et al. [8] stochastically represented the spread process and formulated the pairs trading problem as a stochastic optimal control problem. They then optimized the holding weight by solving the corresponding HJM equation in a closed form in a continuous time setting. This approach was refined by Tourin and Yan [10], who optimized the holding weight when investors make unsymmetrical long/short positions. Yamada and Primbs [12] formulated an optimal pairs trading weight using multiple pairs as a conditional mean-variance model. They solved the model in a closed form and demonstrated its efficiency.

In empirical evaluation, Gatev et al. [4] tested a traditional pairs trading strategy on real stock market data in the United States. They calculated the opening threshold as two times standard deviations from the mean using daily historical price data, and constructed a position when the spread diverges by more than the opening threshold. They backtested the strategy from 1996 to 2002 and showed its efficiency. Various stock markets have been analyzed similarly [1, 5].

Recently, Yamamoto and Hibiki [13] studied an optimal pairs trading strategy using a new approach that is located between theoretical and empirical studies. From a theoretical perspective, they defined a spread process as a stochastic process and generated simulation paths using a Monte Carlo simulation. Next, they developed an optimal pairs trading strategy under actual fund management conditions, such as threshold trading rules, transaction costs, discrete rebalance intervals, and finite investment horizons from an empirical perspective. The optimization problem was formulated as a large-scale simulation based non-continuous optimization problem. Because this problem cannot be solved exactly, Yamamoto and Hibiki [13] employed a heuristic mathematical programming approach called derivative free optimization (DFO), which yielded a nearly exact optimal solution.

However, although the proposed approach optimizes the opening/closing thresholds of a single pair, it cannot handle multiple pairs because the DFO method cannot treat a multitude of variables. In actual investment management, fund managers usually handle 1,000 million yen or more. A single-pair trading strategy incurs problems such as market impact costs, liquidity constraints, and inappropriate fund management, so multiple-pairs trading is required. The optimal solution of single-pair trading strategy might be suboptimal in the multiple-pairs trading strategy adopted by actual fund managers.

In this paper, we attempt to optimize the multiple-pairs trading strategy problem. We formulate this problem as a combination of finding an optimal threshold for each pair and deciding an optimal holding weight. This formulation extends the standard pairs trading strategy of Gatev et al. [4] and is easily understood by investors.

Our paper makes two original contributions to the field:

1. **Consider the optimal pairs trading strategy under actual investment management conditions**
   No previous model has optimized the pairs trading strategy using a threshold rule under actual investment management conditions. However, such a model is most commonly followed by practitioners. This paper formulates the model and discusses the characteristics of the optimal pairs trading strategy.
2. **Propose a DFO-based algorithm for solving large-scale simulation based non-continuous optimization problems**
   The proposed problem is formulated as a large-scale simulation based non-continuous optimization problem. The DFO method efficiently and accurately solves problems with few variables but cannot properly handle 10 or more variables. We resolve this problem by proposing an efficient heuristic algorithm.

This paper is organized as follows. Referring to Mudchanatongsuk et al. [8], we explain our problem in Section 2 and formulate it as a large-scale simulation based non-continuous optimization. Section 3 presents the characteristics of the optimal solutions, revealed through computational experiments, and proposes two algorithms that reduce computational load. Section 4 presents a sensitivity analysis of the model parameters. In Section 5, we propose and validate a heuristic algorithm that improves the efficiency of the DFO method. Concluding remarks are presented in Section 6.

2. Formulation

2.1. Spread process

Following Mudchanatongsuk et al. [8] and Yamamoto and Hibiki [13], we formulate the spread process as a stochastic process. Fund managers consider investing their money in N spreads $X_i, i = 1, 2, \ldots, N$ and a risk-free asset M. Each spread consists of two stocks $A_i, B_i, i = 1, 2, \ldots, N$, giving $2N$ stocks in the investable set. However, the same stocks may be used at different spreads. First, we describe the stock B_i and the risk-free asset M at time t as follows:

$$\begin{aligned} dM(t) &= rM(t)dt, \\ dB_i(t) &= \mu_i dt + \sigma_i dZ_i(t),\ i = 1, 2, \ldots, N, \end{aligned} \tag{2.1}$$

where r is the risk-free rate, μ_i is the expected return of stock B_i, σ_i is the standard deviation of stock B_i, and $Z_i(t)$ denotes standard Brownian motion.

Let X_i be the spread process as a logarithmic difference between stocks A_i and B_i, and we assume that the spread follows an Ornstein-Uhlenbeck process. Then, we have

$$X_i(t) = \ln(A_i(t)) - \ln(B_i(t)),\ i = 1, 2, \ldots, N, \tag{2.2}$$

$$dX_i(t) = \kappa_i(\theta_i - X_i(t))dt + \eta_i dW_i(t),\ i = 1, 2, \ldots, N, \tag{2.3}$$

where θ_i is the mean of spread X_i, κ_i is the speed of mean reversion, η_i is the standard deviation of spread X_i, and $W_i(t)$ is standard Brownian motion. Let ρ_{X_i,X_j}, ρ_{B_i,X_j}, and ρ_{B_i,B_j} denote the correlations among standard Brownian motions $Z_i(t)$ and $W_i(t)$.

$$\begin{aligned}dZ_i(t)\cdot dW_j(t) &= \rho_{B_i,X_j}dt,\ i=1,2,\ldots,N;\ j=1,2,\ldots,N,\\ dZ_i(t)\cdot dZ_j(t) &= \rho_{B_i,B_j}dt,\ i=1,2,\ldots,N;\ j=1,2,\ldots,N,\\ dW_i(t)\cdot dW_j(t) &= \rho_{X_i,X_j}dt,\ i=1,2,\ldots,N;\ j=1,2,\ldots,N.\end{aligned}$$

At time t, the value of portfolio $V(t)$ is then described by

$$dV(t)=V(t)\sum_{i=1}^{N}\left\{h_i^A(t)\frac{dA_i(t)}{A_i(t)}+h_i^B(t)\frac{dB_i(t)}{B_i(t)}+h_i^M(t)\frac{dM(t)}{M(t)}\right\}, \tag{2.4}$$

where h_i^A, h_i^B, and h_i^M are the holding weights of A_i, B_i, and M in the portfolio, respectively. At time 0, $h_i^A(0)=-h_i^B(0)$ and $h_i^M(0)=|h_i^A(0)|$.

2.2. Simulation setting

Next, we explain the simulation setting of our optimization problem. We discretize the stochastic processes of stocks A_i, stocks B_i, and the spreads X_i formulated in Equations (2.1-2.3) as follows:

$$X_i(t_{k+1}) = \theta_i(1-e^{-\kappa_i\Delta t})+e^{-\kappa_i\Delta t}X_i(t_k)+\sqrt{\frac{\eta_i^2}{2\kappa_i}(1-e^{-2\kappa_i\Delta t})}\delta_i(t_k), \tag{2.5}$$

$$B_i(t_{k+1}) = e^{\mu_i\Delta t+\sigma_i\varepsilon_i(t_k)\sqrt{\Delta t}}B_i(t_k), \tag{2.6}$$

$$A_i(t_{k+1}) = B_i(t_{k+1})e^{X_i(t_{k+1})}, \tag{2.7}$$

where $\delta_i(t_k)$ and $\varepsilon_i(t_k)$ are random numbers sampled from the standard normal distribution with the following correlations:

$$\begin{aligned}\mathrm{correl}(\delta_i(t_k),\varepsilon_j(t_k)) &= \rho_{B_i,X_j},\ i=1,2,\ldots,N;\ j=1,2,\ldots,N,\\ \mathrm{correl}(\varepsilon_i(t_k),\varepsilon_j(t_k)) &= \rho_{B_i,B_j},\ i=1,2,\ldots,N;\ j=1,2,\ldots,N,\\ \mathrm{correl}(\delta_i(t_k),\delta_j(t_k)) &= \rho_{X_i,X_j},\ i=1,2,\ldots,N;\ j=1,2,\ldots,N.\end{aligned}$$

We then generate S sample paths by the Monte Carlo method.

At time t_{k+1}, the rate of return of the portfolio on path s is described by

$$\begin{aligned}r_V^{(s)}(t_{k+1}) &= \frac{V^{(s)}(t_{k+1})-V^{(s)}(t_k)}{V^{(s)}(t_k)}\\ &= \sum_{i=1}^{N}\left\{h_{Ai}^{(s)}(t_k)\frac{A_i^{(s)}(t_{k+1})-A_i^{(s)}(t_k)}{A_i^{(s)}(t_k)}+h_{Bi}^{(s)}(t_k)\frac{B_i^{(s)}(t_{k+1})-B_i^{(s)}(t_k)}{B_i^{(s)}(t_k)}+h_{Mi}^{(s)}(t_k)r\Delta t\right\},\end{aligned} \tag{2.8}$$

where $V^{(s)}(t_k)$, $A_i^{(s)}(t_k)$ and $B_i^{(s)}(t_k)$ are values of the portfolio and stocks A_i and B_i respectively, on path s at time t_k. Let $h_{Ai}^{(s)}(t_k)$, $h_{Bi}^{(s)}(t_k)$, and $h_{Mi}^{(s)}(t_k)$ be the weights of stocks A_i, B_i, and the risk-free asset M, respectively.

We then define a transaction cost function $c_V^{(s)}(t_k)$ as a linear function of the portfolio's turnover.

$$c_V^{(s)}(t_k)=c\left(\sum_{i=1}^{N}|h_{Ai}^{(s)}(t_k)-h_{A0i}^{(s)}(t_k)|+|h_{Bi}^{(s)}(t_k)-h_{B0i}^{(s)}(t_k)|\right), \tag{2.9}$$

where c is the coefficient of transaction cost and $h^{(s)}_{A0i}(t_k)$ and $h^{(s)}_{B0i}(t_k)$ are the weights of stocks A_i and B_i, respectively, at the beginning of period t_k. The weights are given by

$$h^{(s)}_{A0i}(t_k) = \frac{h^{(s)}_{Ai}(t_k)\frac{A^{(s)}_i(t_k)}{A^{(s)}_i(t_{k-1})}}{\frac{V^{(s)}(t_k)}{V^{(s)}(t_{k-1})}}, \quad h^{(s)}_{B0i}(t_k) = \frac{h^{(s)}_{Bi}(t_k)\frac{B^{(s)}_i(t_k)}{B^{(s)}_i(t_{k-1})}}{\frac{V^{(s)}(t_k)}{V^{(s)}(t_{k-1})}}.$$

2.3. Objective function

The objective function is defined by using the return of the portfolio based on the weights $h^{(s)}_{Ai}(t_k)$ and $h^{(s)}_{Bi}(t_k)$. Practitioners usually evaluate their strategy by return, risk, and cost. The empirical studies of Gatev et al. [4], Adachi [1], and Hakamada [5] are also based on these components. On path s, the expected return, expected cost, and portfolio risk from t_1 to t_K, respectively, are defined by

$$\begin{aligned} r(s) &= \frac{1}{K}\sum_{k=1}^{K} r^{(s)}_V(t_k), \\ c(s) &= \frac{1}{K}\sum_{k=1}^{K} c^{(s)}_V(t_k), \\ v(s) &= \frac{1}{K}\sum_{k=1}^{K} (r^{(s)}_V(t_k) - r(s))^2. \end{aligned}$$

The objective function is then defined as follows:

$$f = \frac{1}{S}\sum_{s=1}^{S}\{r(s) - \alpha_1 c(s) - \alpha_2 v(s)\}, \tag{2.10}$$

where α_1, and α_2 are the penalties of cost and risk, respectively.

2.4. Transaction rule

The decision variables in our problem are the holding weights $h^{(s)}_{Ai}(t_k)$ and $h^{(s)}_{Bi}(t_k)$. In general, if the holding weights achieve a high objective value, no constraints should be imposed. However, unconstrained holding weights are difficult to manage because they can dynamically change under the high degree of freedom of the problem. Moreover, the optimization problem becomes too complex to solve.

Elliott et al. [3] showed that fund management based on the threshold rule (as explained in Section 1) is an optimal pairs trading strategy. Because it is easily understood, the threshold rule commonly used by investors and by researchers engaged in empirical studies.

Therefore, we apply this rule to decide the holding weights and calculate the optimal thresholds by using the DFO method. First, we define the constant opening/closing thresholds τ^o_i and τ^c_i of spread X_i. At time t_k, the holding weights on path s are decided by the following rules (Figure 2):

Case 1. $X^{(s)}_i(t_k) \geq \theta_i + \tau^o_i \eta_i \sqrt{\Delta t}$ **and** $h^{(s)}_{A0i}(t_k) \geq 0$

$$h^{(s)}_{Ai} = -\{h^{(s)}_{A0i} + h^{(s)}_{B0i} + h^{(s)}_{M0i}\}, \;\; h^{(s)}_{Bi} = h^{(s)}_{A0i} + h^{(s)}_{B0i} + h^{(s)}_{M0i}, \;\; h^{(s)}_{Mi} = 0.$$

Case 2. $\theta_i + \tau^c_i \eta_i \sqrt{\Delta t} > X^{(s)}_i(t_k) \geq \theta_i - \tau^c_i \eta_i \sqrt{\Delta t}$ **and** $h^{(s)}_{A0i}(t_k) \neq 0$

$$h^{(s)}_{Ai} = 0, \;\; h^{(s)}_{Bi} = 0, \;\; h^{(s)}_{Mi} = 0.$$

Case 3. $X_i^{(s)}(t_k) < \theta_i - \tau_i^o \eta_i \sqrt{\Delta t}$ **and** $h_{A0i}^{(s)}(t_k) \leq 0$

$$h_{Ai}^{(s)} = h_{A0i}^{(s)} + h_{B0i}^{(s)} + h_{M0i}^{(s)}, \quad h_{Bi}^{(s)} = -\{h_{A0i}^{(s)} + h_{B0i}^{(s)} + h_{M0i}^{(s)}\}, \quad h_{Mi}^{(s)} = 0.$$

Case 4. Otherwise

$$h_{Ai}^{(s)} = h_{A0i}^{(s)}, \quad h_{Bi}^{(s)} = h_{B0i}^{(s)}, \quad h_{Mi}^{(s)} = h_{M0i}^{(s)}.$$

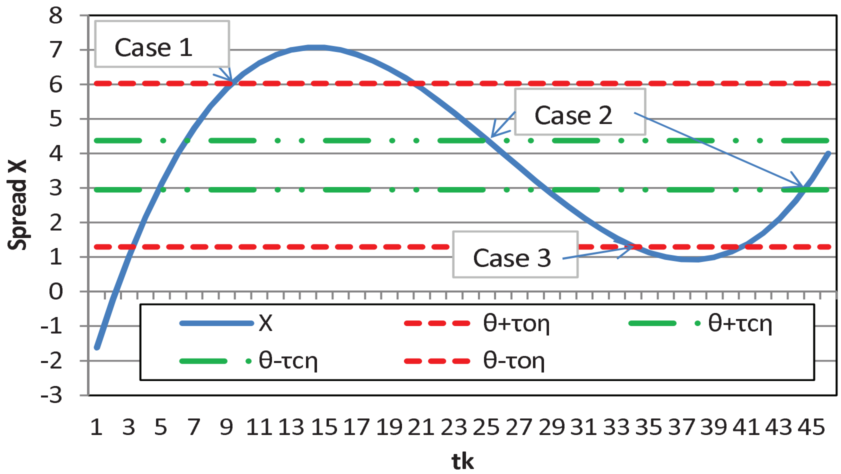

Figure 2: Example of threshold rule (see subsection 2.4 for details)

Yamamoto and Hibiki [13] evaluated two threshold rules: a constant threshold rule and a time-dependent threshold rule. The latter rule generalizes the opening and closing thresholds (τ_i^o and τ_i^c respectively) to $\tau_i^o(t_k)$ and $\tau_i^c(t_k)$, respectively. Yamamoto and Hibiki [13] evaluated $\tau_i^o(t_k)$ and $\tau_i^c(t_k)$ in several test functions and showed that the objective values are almost identical to those of the constant threshold rule.

2.5. Formulation of our optimization problem

Finally we define our optimization problem. To optimize the opening/closing thresholds (τ_i^o and τ_i^c), and the optimal holding weights of each pair, we solve the following problem:

$$\left|\begin{array}{ll} \text{Maximize} & f = \dfrac{1}{S}\displaystyle\sum_{s=1}^{S}\{r(s) - \alpha_1 c(s) - \alpha_2 v(s)\} \\ \text{Subject to} & \tau_i^o \geq \tau_i^c \geq 0, \; i = 1, 2, \ldots, N \\ & h_i \geq 0, \; i = 1, 2, \ldots, N \\ & \displaystyle\sum_{i=1}^{N} h_i = 1, \end{array}\right. \qquad (2.11)$$

where h_i is the initially allocated weight of pair i. On the basis of this weight, investors allocate their money to each pair. Thereafter, they manage each pair by applying the threshold rule to the optimal thresholds τ_i^o and τ_i^c. This large-scale simulation based optimization problem includes non-continuous variables under a predetermined rule. Such a problem cannot be solved by a standard mathematical programming approach but is tractable using the DFO method, a non-linear optimization method that does not require the objective function derivative. As is well known, problems with a convex-minimized objective function and several decision variables find a precise optimal solution (Conn et al. [2]). Hibiki and Yamamoto [6, 7] applied the DFO method in their recent studies of financial applications.

3. Basic Analysis

3.1. Problem setting

In this section, we apply our model to real market data and perform numerical experiments. We selected 10 pairs from different industries and acquired their daily rate of return from April 1, 2013 to March 31, 2014. The pairs are summarized in Table 1.

Table 1: Pair samples selected for the computational study

Pair	Industry	Code	Companies
1	Construction	1925	Daiwa House Industry Co., Ltd.
		1928	Sekisui House, Ltd.
2	Pulp and Paper	3861	Oji Holdings Corporation
		3865	Hokuetsu Kishu Paper Co., Ltd
3	Machinery	6301	Komatsu Ltd.
		6305	Hitachi Construction Machinery Co., Ltd.
4	Transportation Equipment	7203	Toyota Motor Corporation
		6201	Toyota Industries Corporation
5	Electric Power and Gas	9531	Tokyo Gas Co., Ltd.
		9532	Osaka Gas Co., Ltd.
6	Air Transportation	9201	Japan Airlines Co., Ltd
		9202	ANA Holdings Inc.
7	Wholesale Trade	8058	Mitsubishi Corporation
		8002	Marubeni Corporation
8	Retail Trade	3382	Seven & I Holdings Co., Ltd.
		8278	Fuji Co., Ltd.
9	Banks	8306	Mitsubishi UFJ Financial Group, Inc.
		8316	Sumitomo Mitsui Financial Group, Inc.
10	Real Estate	8802	Mitsubishi Estate Company, Limited
		8830	Sumitomo Realty & Development Co., Ltd.

Following Yamamoto and Hibiki [13], we estimated the simulation parameters of each pair from the daily rate of return. The results are shown in Table 2*.

Other parameters were set as follows: the number of simulation $S = 10,000$, period $T = 250$, rebalance interval $\Delta t = 0.004(1/250)$, risk-free rate $r_f = 0.04\%$, and transaction cost coefficient $c = 0.3\%$. The problem was solved on Panasonic Let's note(Windows 7, 2.53GHz, 4GB) using Numerical Optimizer Ver.16.1(add on Numerical Optimizer/DFO) (NTT DATA Mathematical Systems Inc. [9]).

3.2. Problem characteristics

In this subsection, we characterize the optimal solution of the two pairs problem. The input data are pairs 1 and 2 in Table 1.

The results for $\alpha_1 = 1$, $\alpha_2 = 1$ are presented in Table 3. Brackets show the optimal opening/closing thresholds of the equivalent single-pair problem ($N = 1$) proposed by Yamamoto and Hibiki [13].

As shown in the table, the problem was solved in approximately 1,300 seconds, and the optimal opening/closing thresholds are consistent with the single-pair problem. Moreover, the optimal holding weights differ from the 50%($1/N$ portfolio) used in early studies.

*Correlation coefficients among all pairs are omitted because of space limitations.

Table 2: Estimated parameters

(a) Spread parameters

Pair	μ_i	σ_i	κ_i	θ_i	η_i
1	0.31	0.22	12.21	0.38	0.18
2	0.07	0.27	11.25	-0.12	0.19
3	0.07	0.28	9.94	0.12	0.16
4	0.35	0.24	18.13	0.19	0.15
5	0.26	0.18	7.18	0.37	0.14
6	0.36	0.21	16.90	2.46	0.19
7	0.01	0.20	0.24	2.03	0.14
8	0.06	0.22	1.08	0.91	0.19
9	0.04	0.23	4.03	-1.86	0.14
10	0.07	0.29	4.53	-0.48	0.17

(b) Correlations between pairs 1 and 2

	B_1	X_1	B_2	X_2
B_1	1.00	-0.25	0.44	0.07
X_1		1.00	0.12	-0.05
B_2			1.00	-0.50
X_2				1.00

Table 3: Computational results of the two-pairs problem

CPU(sec)	Obj. Val.	h_1	h_2	τ_1^o	τ_2^o	τ_1^c	τ_2^c
1386.09	0.194	35.0%	65.0%	2.629	2.788	0.001	0.001
				(2.605)	(2.786)	(0.001)	(0.000)

Next, we verify the shape of the objective function. Setting τ_2^o, τ_1^c and τ_2^c to the optimal values in Table 3, we have two decision variables τ_1^o and h_1. The left and right panels of Figure 3 show how the objective value depends on the holding weight h_1 (with $\tau_1^o = 2.629$) and the opening threshold τ_1^o (with $h_1 = 35\%$), respectively.

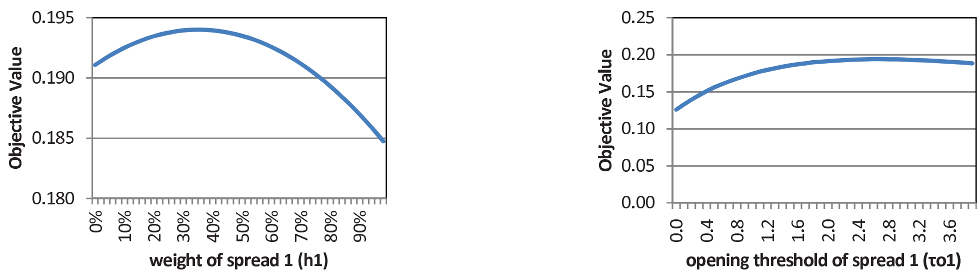

Figure 3: Objective value as a function of holding weight (left) and opening threshold (right)

Both plots are concave, confirming that the DFO method will likely find the global optimal solution.

3.3. Contribution of decision variables to objective value

In this subsection, we computationally determine the contributions of the decision variables to the objective value. The four optimization problems are defined in Table 4.

Strategy 1: Same threshold and equal allocation

This simple strategy is adopted in almost all empirical studies. Following Gatev et al. [4], we set the thresholds $\tau_i^o = 2.0$, $i = 1, 2, \ldots, N$ and $\tau_i^c = 0.0$, $i = 1, 2, \ldots, N$ and the holding weight $h_i = 1/N$, $i = 1, 2, \ldots, N$.

Table 4: Summary of evaluated strategies

	Equal Allocation	Optimal Allocation
Same Threshold	Strategy 1	Strategy 3
Different Thresholds	Strategy 2	Strategy 4

Strategy 2: Different thresholds and equal allocation
We use the optimal thresholds for each pair to solve the single-pair trading model proposed by Yamamoto and Hibiki [13] and the same holding weights $h_i = 1/N$, $i = 1, 2, \ldots, N$. This strategy is an extension of Strategy 1.

Strategy 3: Same threshold and optimal allocation
This strategy solves the problem (2.11), fixing the opening/closing thresholds to those used in Strategy 1 and optimizing the holding weight. This strategy is another extension of Strategy 1.

Strategy 4: Different thresholds and optimal allocation
This strategy implements our proposed model, solving the problem (2.11) by the DFO method.

The optimal solutions of each strategy are listed in Table 5. Underlines show the decision variables.

Table 5: Comparisons of the evaluation strategies

Strategy	Obj. Val.	h_1	h_2	τ_1^o	τ_2^o	τ_1^c	τ_2^c
Strategy 1	0.1864	50.0%	50.0%	2.000	2.000	0.000	0.000
Strategy 2	0.1936	50.0%	50.0%	<u>2.605</u>	<u>2.786</u>	<u>0.001</u>	<u>0.000</u>
Strategy 3	0.1868	<u>37.2%</u>	<u>62.8%</u>	2.000	2.000	0.000	0.000
Strategy 4	0.1940	<u>35.0%</u>	<u>65.0%</u>	<u>2.628</u>	<u>2.788</u>	<u>0.001</u>	<u>0.001</u>

Relative to Strategy 1, Strategy 2 increases the objective value by 3%, but Strategy 3 exerts little effect. This indicates that the objective value benefits from varying the opening threshold but is relatively insensitive to the holding weight, probably because the optimal holding weight is close to 50%. The optimal solutions of Strategies 2 and 4 are almost identical, indicating that the holding weight and opening threshold are independent variables.

3.4. Separation of the problem

Problem (2.11) (Strategy 4) is intractable when the number of pairs N is large, because the DFO method is unsuitable for problems with many decision variables. On the basis of the above results, we propose two algorithms for solving the problem (2.11).

Separated Problem 1:SP1

Step 1. Problem (2.11) is solved by the DFO method under the following constraints:

$$h_i = 1/N,\ \tau_i^c = 0,\ i = 1, 2, \ldots, N.$$

Step 2. Let $(\tau_i^o)^*$, $i = 1, 2, \ldots, N$ be the optimal solution of Step 1. The optimal holding weight is derived by the problem (2.11) under the following constraints:

$$\tau_i^o = (\tau_i^o)^*,\ \tau_i^c = 0,\ i = 1, 2, \ldots, N.$$

Separated Problem 2:SP2

Step 1. Optimize the opening/closing thresholds τ_i^o and τ_i^c by solving the problem (2.11) for each pair with $N = 1$, as proposed by Yamamoto and Hibiki [13]. Then we find the optimal opening/closing thresholds $(\tau_i^o)^*$ and $(\tau_i^c)^*$.

Step 2. The optimal holding weight is derived by the problem (2.11) under the following constraints:

$$\tau_i^o = (\tau_i^o)^*,\ \tau_i^c = (\tau_i^c)^*,\ i = 1, 2, \ldots, N.$$

The SP1 formulation assumes that the holding weight and thresholds do not interact, as inferred in subsection 3.3. Meanwhile, the SP2 formulation is based on the concordance of the thresholds of problem (2.11) in the cases of $N > 1$ and $N = 1$, as inferred in subsection 3.2.

Step 1 of SP1 optimizes N decision variables in one run of the DFO method. On the other hand, Step 1 of SP2 optimizes two decision variables in N runs of the DFO method. The computational time of this problem should rapidly increase with increasing number of decision variables. Thus, the computational time of SP2 is smaller than that of SP1 when the number of decision variables is large.

The optimal solutions of each problem are compared in Table 6. Hereafter, the optimization problem (2.11) is called "Integrated Problem (IP)".

Table 6: Optimal solutions to each problem

Problem	CPU time(sec)	Obj. Val.	h_1	h_2	τ_1^o	τ_2^o	τ_1^c	τ_2^c
IP	1386.0	0.19404	35.0%	65.0%	2.629	2.788	0.001	0.001
SP1	324.8	0.19400	35.3%	64.7%	2.610	2.799	0.000	0.000
SP2	697.9	0.19406	35.4%	64.6%	2.605	2.786	0.001	0.000

The three methods yield very similar objective values. However, the computational time of IP is approximately double that of SP2 and quadruple that of SP1. This confirms that the holding weight and threshold are independent variables. Moreover, the optimal thresholds of problem (2.11) are almost identical for $N > 1$ and $N = 1$.

Table 7 lists the objective values and computational times for different numbers of pairs N. This computation uses the first N pairs in Table 1. The reported results are averaged over 10 cases with different random seeds. The standard deviations of the CPU times are indicated in parentheses. The upper limit of the computation time was set to 30,000 seconds.

When $N = 5$, the IP requires 20,000 seconds of runtime versus approximately 5,000 and 4,000 seconds for SP1 and SP2, respectively. All approaches yield the same objective value. Therefore, the proposed algorithms (SP1 and SP2) find the same optimal solution as IP within a much shorter time.

When $N \geq 4$, the computational time of SP1 exceeds that of SP2. Above, we mentioned that the computational time of Step 1 in SP1 grows rapidly with N. Thus, SP2 is more effective than SP1 when the number of pairs N is large.

4. Sensitivity Analysis

This section clarifies the characteristics of the optimal pairs trading strategy and checks the accuracy of the optimal solution to the separated problem. To this end, we conducted

Table 7: Performances of the evaluated methods for different number of pairs N

	N	2	3	4	5	6	7	8	9	10
Obj.	IP	0.1935	0.1932	0.1937	0.1935	-	-	-	-	-
Val.	SP1	0.1935	0.1932	0.1937	0.1934	0.2258	0.2266	0.2268	0.2265	-
	SP2	0.1935	0.1932	0.1936	0.1934	0.2258	0.2266	0.2268	0.2265	0.2262
CPU	IP	1346	3857	9729	19194	-	-	-	-	-
time		(229)	(1726)	(2656)	(5319)	-	-	-	-	-
(sec)	SP1	376	1129	2548	5139	9183	13202	21827	27847	-
		(43)	(207)	(805)	(1040)	(1237)	(3991)	(4428)	(4849)	-
	SP2	535	1237	2217	3599	4733	6879	8892	11655	14475
		(86)	(206)	(218)	(373)	(252)	(331)	(850)	(945)	(1357)

a sensitivity analysis of the model parameters in the two-pairs problem (pairs 1 and 2 in Table 1). We report the averages of 10 cases evolved from different random seeds.

First, we confirm the sensitivity of the optimal solution to the spread parameter. Figure 4 plots the objective values of the integrated and second separated problems (IP and SP2) and the relative error of these objective values ((IP−SP2)/SP2), respectively (left panel) and the optimal opening thresholds of the two pairs and the holding weight of pair 1 (right panel) as functions of the speed of spread 1. The optimal closing threshold is almost zero in all cases.

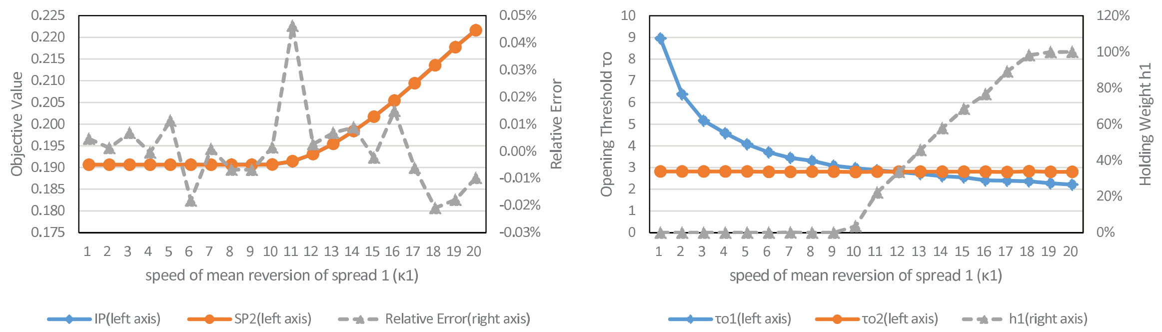

Figure 4: Optimization results as functions of spread speed of pair 1 κ_1

The objective values of IP and SP2 are almost identical. As the spread speed of pair 1 increases, the objective value increases, the opening threshold of pair 1 decreases, and the holding weight of pair 1 increases. This means that the number of constructing a position by setting a small opening threshold increases, enabling a high return.

Figure 5 relates the objective value and optimal solution to the volatility of spread 1 (η_1).

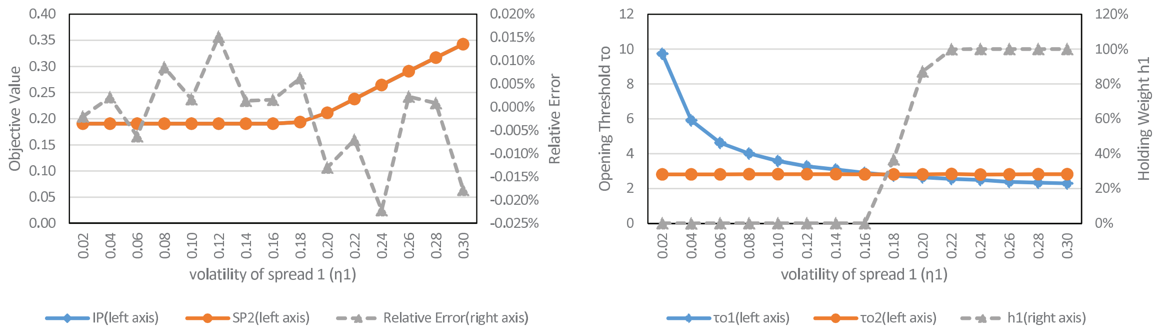

Figure 5: Optimization results as functions of volatility in the spread of pair 1 η_1

As a function of the volatility of spread 1, the objective value and optimal solution behave similarly to Figure 4, indicating that spread speed and volatility exert similar effects on the return of pair 1. The same results were reported for $N = 1$ in Yamamoto and Hibiki [13].

Next, we relate the optimal solution to the correlations between pairs 1 and 2, specifically, the correlations between spreads X_1 and X_2, spread X_1 and stock B_2, and stocks B_1 and B_2.

The relationship between the optimal solution and the correlation between spreads X_1 and X_2 (ρ_{X_1,X_2}) in plotted in Figure 6.

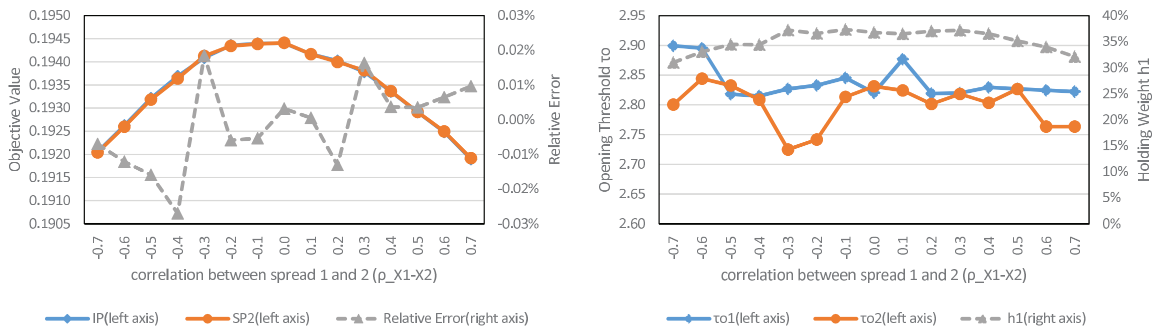

Figure 6: Optimization results as functions of the spread correlation ρ_{X_1,X_2}

The objective value is maximized at correlations around zero, thus, the optimal holding weight is approximately 50%. This means that diversification affects the objective value by decreasing the portfolio risk.

Figures 7 and 8 plot the results versus the correlations between spread X_1 and stock B_2 (ρ_{X_1,B_2}), and between stocks B_1 and B_2 (ρ_{B_1,B_2}), respectively.

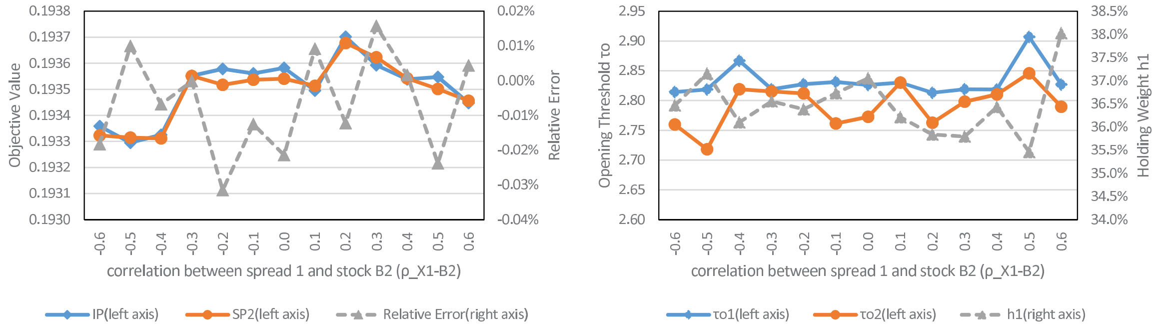

Figure 7: Optimization results as functions of the spread X_1-stock B_1 correlation ρ_{X_1,B_2}

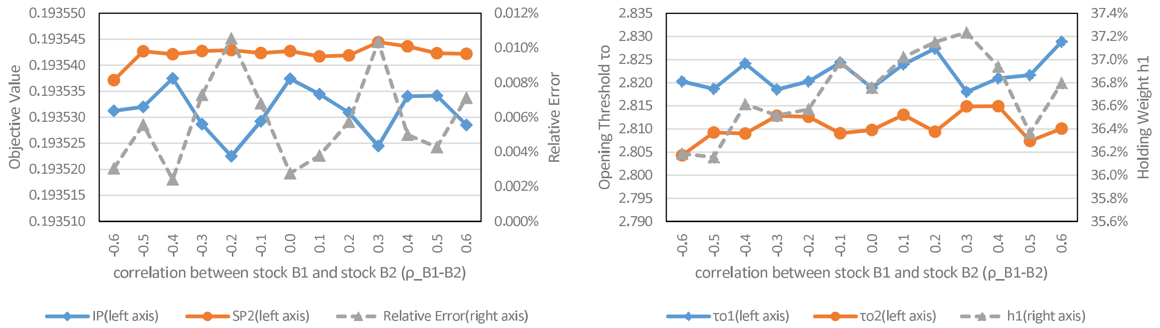

Figure 8: Optimization results as functions of the stock correlation ρ_{B_1,B_2}

These figures confirm that the optimal solution is much less sensitive to ρ_{X_1,B_2} and ρ_{B_1,B_2}. In addition, the objective values of IP and SP2 coincide in almost all cases.

Finally, we relate the optimal solution to the penalty parameters. The relationships between the optimization results and the cost and risk penalties (α_1, α_2, respectively) are plotted in Figures 9 and 10, respectively.

As the penalties increase, the optimal solutions disfavor construction of a position, as determined for $N = 1$ in Yamamoto and Hibiki [13]. From these results, The proposed algorithm (SP2) finds the same optimal solution as the integrated problem (IP) within a shorter computational time.

Finally, we summarize the characteristics of the optimal pairs trading strategy determined in the above analysis.

- Almost all closing thresholds are zero because the spread process was assumed as a mean reversion process.
- The holding weight is independent of the opening/closing thresholds.
- The opening thresholds do not interact in the multiple-pairs analysis. Thus the single-pair and multiple-pairs problem yield the same opening thresholds.
- The opening threshold is affected by the speed and volatility of the spread of a pair.

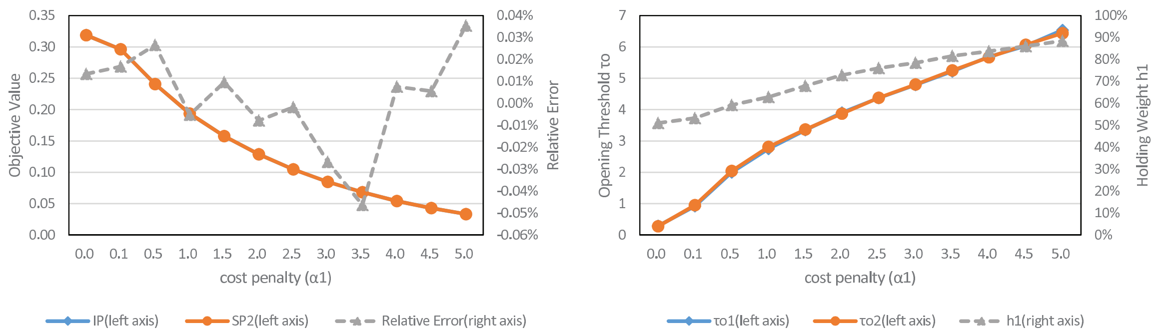

Figure 9: Optimization results as functions of cost penalty α_1

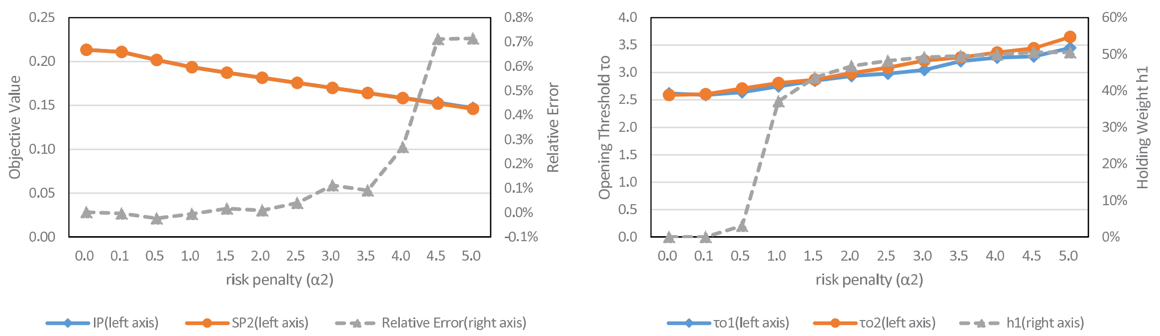

Figure 10: Optimization results as functions of risk penalty α_2

- The objective value is high when two spreads are completely uncorrelated but is unaffected by the spread-stock and stock-stock correlations.
- In the optimal solution, the investor should not construct a position when the risk and cost penalties are high.

5. Algorithm for Solving Large-scale Problems

5.1. Algorithm

This section proposes our efficient heuristic algorithm for solving large-scale problems. This algorithm is based on the characteristics of the optimal solution in the former analysis.

First, we add the following four settings to the model.

1. Fix the closing thresholds to zero

In Sections 3 and 4 of the present paper and also in Yamamoto and Hibiki [13], the closing thresholds were almost always zero. Moreover, a zero closing threshold was assumed in early empirical studies such as Gatev et al. [4]. Therefore, we set

$$\tau_i^c = 0,\ i = 1, 2, \ldots, N.$$

2. Represent the opening thresholds by an exponential function

In Section 4 and also in Yamamoto and Hibiki [13], the opening thresholds decreased with increasing volatility of the spread. Therefore, we express the opening thresholds by the following exponential function of the volatility:

$$\tau_i^o = p_1(1 + p_2 e^{-p_3 \eta_i}),\ i = 1, 2, \ldots, N,$$

where p_1, p_2, and p_3 are parameters of the exponential function. Under this assumption, the number of decision variables is three for any number of pairs.

3. Separate the optimal threshold and optimal allocation problems

To reduce the computational time, we separate the optimal threshold and optimal allocation problems.

4. Optimize the holding weight by a mean-variance model

To further reduce the computational time, we replace the simulation based optimization model with a standard mean-variance model to optimize the holding weight.

Then we propose the following heuristic algorithm by iteratively solving the separated problems.

Algorithm. Iterative Method

Step 1. Initialization.

$k = 1,\ \hat{h}_{i1} = 1/N,\ i = 1, 2, \ldots, N,\ f_0 = -\infty,\ \epsilon = 0.1\%.$

Step 2. Solve the optimal threshold problem.

Solve the following optimization problem by the DFO method:

$$\left|\begin{array}{ll} \text{Maximize} & f = \dfrac{1}{S}\displaystyle\sum_{s=1}^{S}\{r(s) - \alpha_1 c(s) - \alpha_2 v(s)\} \\ \text{Subject to} & \tau_i^o = p_1(1 + p_2 e^{-p_3 \eta_i}),\ i = 1, 2, \ldots, N \\ & \tau_i^c = 0,\ i = 1, 2, \ldots, N \\ & h_i = \hat{h}_{ik}\ i = 1, 2, \ldots, N \\ & p_1 \geq 0,\ p_2 \geq 0,\ p_3 \geq 0. \end{array}\right. \tag{5.1}$$

This step solves the problem (2.11) with fixed holding weights and closing thresholds. It calculates the optimal parameters p_1^*, p_2^*, and p_3^*, denoting the decision variables of the exponential function.

After solving the problem (5.1), the optimal thresholds are calculated from the optimal parameters p_1^*, p_2^*, and p_3^* as follows:

$$\begin{aligned} (\tau_i^o)^* &= p_1^*(1 + p_2^* e^{-p_3^* \eta_i}),\ i = 1, 2, \ldots, N, \\ (\tau_i^c)^* &= 0,\ i = 1, 2, \ldots, N. \end{aligned}$$

Step 3. Solve the optimal allocation problem.

Let $m_i^{(s)}$ be the expected return, $c_i^{(s)}$ be the transaction cost of pair i and $v_{ij}^{(s)}$ be the covariance of pairs i and j on path s. These parameters are determined in the simulation using the calculated thresholds $(\tau_i^o)^*$ and $(\tau_i^c)^*$ and are averaged as follows:

$$\begin{aligned} m_i &= \frac{1}{S}\sum_{s=1}^{S} m_i^{(s)},\ i = 1,2,\ldots,N, \\ c_i &= \frac{1}{S}\sum_{s=1}^{S} c_i^{(s)},\ i = 1,2,\ldots,N, \\ v_{ij} &= \frac{1}{S}\sum_{s=1}^{S} v_{ij}^{(s)},\ i = 1,2,\ldots,N;\ j = 1,2,\ldots,N. \end{aligned}$$

Then the optimal allocation h_i^*, $i = 1,2,\ldots,N$ is determined by solving the following quadratic programming problem:

$$\left|\begin{array}{ll} \text{Maximize} & \sum_{i=1}^{N}(m_i - \alpha_1 c_i)h_i - \alpha_2 \sum_{i=1}^{N}\sum_{j=1}^{N} h_i h_j v_{ij} \\ \text{Subject to} & \sum_{i=1}^{N} h_i = 1 \\ & h_i \geq 0,\ i = 1,2,\ldots,N. \end{array}\right. \tag{5.2}$$

Step 4. Evaluate the objective function.

Using the obtained $(\tau_i^o)^*$, $(\tau_i^c)^*$ and h_i^*, calculate the objective value f_k by the problem (2.11). The algorithm terminates when $(f_k - f_{k-1})/f_{k-1} \leq \epsilon$. Otherwise, it sets $\hat{h}_{ik+1} = h_i^*$, $i = 1,2,\ldots,N$, $k \leftarrow k+1$, and returns to Step 2.

5.2. Computational experiments

To validate our proposed heuristic algorithm, we examine the relationship between the optimal solution and the number of pairs. For this purpose, we solve the same problems in Table 7 using proposed heuristic algorithm. The results are appended to the results of Table 7 in Table 8. Here, the heuristic algorithm is denoted by “HA”.

Table 8: Performances of the evaluated methods for different number of pairs N

	N	2	3	4	5	6	7	8	9	10
Obj.	IP	0.1935	0.1932	0.1937	0.1935	-	-	-	-	-
Val.	SP1	0.1935	0.1932	0.1937	0.1934	0.2258	0.2266	0.2268	0.2265	-
	SP2	0.1935	0.1932	0.1936	0.1934	0.2258	0.2266	0.2268	0.2265	0.2262
	HA	0.1935	0.1932	0.1936	0.1934	0.2255	0.2265	0.2267	0.2264	0.2259
CPU	IP	1346	3857	9729	19194	-	-	-	-	-
time		(229)	(1726)	(2656)	(5319)	-	-	-	-	-
(sec)	SP1	376	1129	2548	5139	9183	13202	21827	27847	-
		(43)	(207)	(805)	(1040)	(1237)	(3991)	(4428)	(4849)	-
	SP2	535	1237	2217	3599	4733	6879	8892	11655	14475
		(86)	(206)	(218)	(373)	(252)	(331)	(850)	(945)	(1357)
	HA	1124	1848	2501	3348	5939	5140	5701	6739	7424
		(247)	(305)	(309)	(1350)	(1342)	(693)	(492)	(1059)	(626)

The heuristic algorithm significantly reduces the computational time when the number of pairs is large ($N \geq 7$). In particular, when $N = 9$ or 10, the heuristic algorithm returns an optimal solution in approximately half the time of SP2. Moreover, all algorithms return very similar objective values.

These results confirm that our proposed heuristic algorithm efficiently solves the target problem when the number of pairs is large without compromising the computational accuracy.

6. Conclusions and Future Directions

We formulated the optimal pairs trading strategy problem for multiple pairs under actual fund management conditions. The threshold trading rule and other problem settings refer to practical fund management and empirical studies. When the number of pairs is small, this problem can be accurately solved by the derivative free optimization (DFO) method. In a series of computational experiments, we determined the characteristics of the solutions.

From the results of the computation experiments, we also proposed an efficient algorithm for solving this optimization problem on large scales. In computational validation tests, the heuristic algorithm solved the problem within a practical computational time while maintaining the accuracy of the objective value.

We believe that our model will assist practitioners adopting this investment strategy and application studies based on the DFO method.

Acknowledgments

This work was supported by Grant-in-Aid for Scientific Research (C) 15K01201 and Grant-in-Aid for Young Scientists (B) 16K16359.

References

[1] T. Adachi: Analysis of the profitability of pairs trading strategy. *Proceeding of Society of Applied Economic Time Series Analysis* (2006), 25–50 (in Japanese).

[2] A.R. Conn, K. Scheinberg, and W.P. Malcom: *Introduction to Derivative-Free Optimization* (SIAM, Philadelphia, 2009).

[3] R.J. Elliott, J.V.D. Hoek, and W.P. Malcom: Pairs trading. *Quantitative Finance*, **5** (2005), 271–276.

[4] E.G. Gatev, W.N. Goetzmann, and K.G. Rouwenhorst: Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, **19** (2006), 797–827.

[5] M. Hakamada: Study of the profitability of the pairs trading strategy in the Japanese stock market. *Proceeding of Nippon Finance Association* (2002), 51–57 (in Japanese).

[6] N. Hibiki, R. Yamamoto, T. Tanabe, and Y. Imai: Optimal asset allocation strategy with transaction costs —derivation of the optimal no-trade region using derivative free optimization—. *Transactions of the Operations Research Society of Japan*, **57** (2013), 1–26 (in Japanese).

[7] N. Hibiki and R. Yamamoto: Optimal symmetric no-trade ranges in asset rebalancing strategy with transaction costs. *Asia-Pacific Journal of Risk and Insurance*, **8** (2014), 293–327.

[8] S. Mudchanatongsuk, J.A. Primbs, and W. Wong: Optimal pairs trading: A stochastic control approach. *Proceedings of the American Control Conference* (2008), 1035–1039.

[9] NTT DATA Mathematical Systems Inc.: *Numerical Optimizer/DFO* (2013) (in Japanese).

[10] A. Tourin and R. Yan: Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control*, **37** (2013), 1972–1981.

[11] G. Vidyamurthy: *Pairs Trading: Quantitative Methods and Analysis* (John Wiley and Sons, New York, 2004).

[12] Y. Yamada and J.A. Primbs: Model predictive control for optimal portfolios with cointegrated pairs of stocks. *Proceeding of 51st IEEE Conference on Decision and Control* (2012), 5705–5710.

[13] R. Yamamoto and N. Hibiki: Optimal pairs trading strategy under actual fund management conditions using derivative free optimization approach. *JAFEE Journal*, (2015), 202–233 (in Japanese).

Rei Yamamoto
Faculty of Economics
Department of Finance
Musashi University
1-26-1, Toyotama-kami, Nerima-ku
Tokyo 176-8534, Japan
E-mail: rei.yamamoto@cc.musashi.ac.jp