Operations Research Letters 46 (2018) 225–232

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Robust dynamic pairs trading with cointegration

Mei Choi Chiu [a,*], Hoi Ying Wong [b]

[a] Department of Mathematics and Information Technology, Education University of Hong Kong, Tai Po, Hong Kong
[b] Department of Statistics, The Chinese University of Hong Kong, Shatin, Hong Kong

ARTICLE INFO

Article history:
Received 26 December 2016
Received in revised form 18 January 2018
Accepted 18 January 2018
Available online 31 January 2018

Keywords:
Cointegration
Pairs trading
Robust decision rule

ABSTRACT

This paper investigates the robust optimal pairs trading using the concept of equivalent probability measures and a penalty function associated with the confidence in parameter estimates when the parameters in the drift term of the continuous-time cointegration model are estimated with errors. A closed-form solution is derived for the robust pairs trading rule. We compare the robust pairs trading rule against its non-robust counterpart using simulations and real data. The robust strategy is empirically more stable and less volatile.

1. Introduction

Pairs trading is an appealing statistical arbitrage strategy in the practical financial market. It attempts to take profit from short-term departure of financial assets which share a common long-term equilibrium. A popular way to identify the relevant assets and the long-term equilibrium uses the concept of cointegration that originated in [10]. Using this 2003 Nobel prize framework, the studies in [1] and [2] pioneer the rigorous investigation of cointegration techniques in active portfolio management. Other applications of cointegration in asset pricing include derivative pricing and commodity pricing in [9] and [8].

The analytic tractability of the continuous-time cointegration model facilitates the theoretical development of optimal pairs trading strategies. The mean–variance pairs trading rules are derived for banks [4], insurance companies [5] and firms with liability [6]. The optimal pairs trading in [18] is based on maximizing the expected constant relative risk aversion (CRRA) utility. The study of [24] considers the expected constant absolute risk aversion (CARA) utility. Cointegration is also useful for longevity risk management using the optimal portfolio approach as shown in [26] and [27]. Kwok et al. [15] apply cointegration to longevity security demand under relative performance concerns. Chiu and Wong [7] mathematically prove the statistical arbitrage feature of cointegration market in the sense of [13], once all of the model parameters in the cointegration model are perfectly known in advance.

The actual profitability of pairs trading is, however, subject to many practical issues. Lei and Xu [17] show that the statistical arbitrage through pairs trading could be costly. They numerically characterize the trading and non-trading regions for a pair of assets with a specific cointegration structure subject to proportional transaction costs and demonstrate empirically that the profit is much reduced for the existence of transaction costs. Another practical issue is the estimation errors of parameters in the cointegration model. The finance literature has reported the distortion of empirical optimal portfolio performance by estimation errors. A remedy to reduce the effect of parameter uncertainty uses the robust policy of [14] and [19], so that investment decision is made under the worst case among scenarios characterized by a set of equivalent probability measures. Therefore, the classic expected utility maximization is replaced by a maximin problem. Although such an approach has been applied to diffusion processes with deterministic drifts ([28] and [22]), the corresponding theoretical treatment for the cointegration is yet to be considered.

Recently, the uncertainty of the mean reversion level in cointegration for portfolio optimization is considered in [16] using the formulation of partial information. Ours using the framework of [19] is significantly different from [16]. The authors of [16] consider the worst-case scenario over an interval for the mean reversion level, but we consider a set of equivalent probability measures and a penalty function associated with the confidence on parameter estimates. The empirical use of our framework can be realized through data-driven cross-validation or machine learning techniques in selecting the penalty function. The partial information setting does not have such a feature. In addition, we extend the framework of [20] on zero-sum stochastic differential games to establish a verification theorem for the approach of [19] under cointegration models.

The remainder of the paper is organized as follows. Section 2 reviews the cointegration model and introduces the robust pairs

* Corresponding author.
E-mail addresses: mcchiu@eduhk.hk (M.C. Chiu), hywong@cuhk.edu.hk (H.Y. Wong).

https://doi.org/10.1016/j.orl.2018.01.006

trading problem with cointegration subject to parameter estimation errors. Section 3 derives the robust optimal solution for CARA investors whose utility functions are exponential functions.[1] Section 4 compares the features of the classical and the robust solutions subject to estimation error through a series of simulations. The comparison is based on various performance measures such as profitability, risk and objective function value. Section 5 offers an empirical study to show the advantage of the robust approach with real data, Concluding remarks are made in Section 6.

2. Problem formulation

2.1. Cointegration model

By Granger's representation theorem [12], the cointegrated vector time series can be expressed as an error correction model (ECM). In a discrete-time model with constant parameters, an error correction dynamic for the n-component asset price time series with k ($1 \leq k \leq n$) cointegrating factors is defined as follows: $\ln S_{i,t} - \ln S_{i,t-1} = \mu_i + \sum_{j=1}^{k}\delta_{ij}z_{j,t-1} + \sum_{j=1}^{n}\sigma_{ij,t}\epsilon_{j,t}$ with $i = 1, \ldots, n$, where $z_{j,t} = a_j + b_j t + \sum_{i=1}^{n} c_{ij}\ln S_{i,t}$ for $j = 1, \ldots, k$, $S_{i,t}$ is the price of asset i at time t for $i = 1, \ldots, n$, $(c_{1j}, \ldots, c_{nj})$ are linearly independent vectors for $j = 1, \ldots, k$, and $[\epsilon_{1,t}, \ldots, \epsilon_{n,t}]$ is a vector of independent identically distributed (i.i.d.) standard normal random variables. In the error correction model, the vector of k cointegrating factors, $[z_{1,t}, \ldots, z_{k,t}]$, should be a stationary time series such that each $z_{j,t}$ has a bounded variance at all time points.

To better motivate our problem formulation, consider two cointegrated assets with one cointegrating factor. The estimation of ECM involves the ordinary least squares (OLS) method that estimates the parameters a, b and c through $\ln S_{1,t} = a+bt+c\ln S_{2,t} - z_t$, where z_t is the noise term in the OLS. If the time series of $\{z_t\}$ is tested to be stationary, then $z_t = a + bt + c\ln S_{2,t} - \ln S_{1,t}$ is the cointegration relationship. After the estimation, a simple strategy constructs a long-short portfolio based on $\widehat{c}\ln S_{2,t} - \ln S_{1,t}$ and determines the optimal allocation between the portfolio and the bank account based on the departure level to its long-term equilibrium trend which is related to $\widehat{b}$, $\widehat{c}$ and the volatility of z_t, see [25].

Stock [23] shows that the OLS estimate of $\widehat{c}$ is super consistent in the sense that it converges to the true value in an order of $\frac{1}{N}$ whereas other estimators $\widehat{a}$ and $\widehat{b}$ are only consistent estimators converging in an order of $\frac{1}{\sqrt{N}}$ for a sample size N. In other words, the estimation error for $\widehat{c}$ is not serious compared to those of $\widehat{a}$ and $\widehat{b}$. Therefore, this study focuses on reducing the effect of estimation error in μ in the general ECM in $\ln S_{1,t} = a + bt + c\ln S_{2,t} - z_t$,. To gain analytic tractability, we use a continuous-time ECM.

2.2. Continuous-time model

In a continuous-time economy, the cointegration model is the diffusion limit of the discrete-time ECM [9]:

$$d\ln S_{i,t} = \left(\mu + \sum_{j=1}^{k}\delta_{ij}z_{j,t}\right)dt + \sum_{j=1}^{n}\sigma_{ij}dW_{j,t}, \quad i = 1, \ldots, n, \quad (1)$$

$$z_{j,t} = a_j + b_j t + \sum_{i=1}^{n} c_{ij}\ln S_{i,t} \quad j = 1, \ldots, k, \quad (2)$$

where $W_{i,t}$, $i = 1, \ldots, n$ are independent standard Wiener processes. To reflect the spirit of the error correction model in its diffusion limit, when $k = 1$, $z_{1,t}$ follows the Ornstein–Uhlenbeck process [9], $dz_{1,t} = \left(\mu_1 - \alpha z_{1,t}\right)dt + \sigma_{z,1}\,d\widehat{W}_{1,t}$, where μ_1 is a real constant, and α and σ_1 are positive constants. In general, the vector of k cointegrating factors, $z_t = [z_{1,t}, \ldots, z_{k,t}]$, satisfies the stochastic differential equation (SDE): $dz_t = (\mu_z - J\cdot z_t)\,dt + \sigma_z d\widehat{W}_t$, where J is a $k \times k$ positive diagonal Jordan matrix for a cointegration system, $\sigma_z\sigma_z'$ is a positive definite matrix of size k, μ_z is a k dimensional vector, and $\widehat{W}_t$ is a vector of independent Wiener processes induced from (2). Specifically, $\mu_z = b + C\mu$, $J = C\delta$ and $C\sigma dW = \sigma_z d\widehat{W}_t$ in distribution, where $b = [b_j] \in \mathbb{R}^k$, $\mu = [\mu_j] \in \mathbb{R}^n$, $C = [c_{ij}] \in \mathbb{R}^{k\times n}$, $\sigma = [\sigma_{ij}] \in \mathbb{R}^{n\times n}$, $\sigma_z \in \mathbb{R}^{k\times k}$ and $\delta = [\delta_{ij}] \in \mathbb{R}^{n\times k}$.

To simplify the notation, we substitute (2) into (1) and write $d\ln S_t = (\theta - A\cdot\ln S_t)\,dt + \sigma dW_t$, where $\ln S_t$ is a vector that contains the log prices of n assets, θ is an n-dimensional vector, W_t is a vector of uncorrelated Wiener processes, $\sigma\sigma'$ represents the $n \times n$ variance–covariance matrix, and A is $n \times n$ coefficient matrix of cointegration. Specifically, $\theta = \mu + \delta(a + bt) \in \mathbb{R}^n$, $A = \delta C \in \mathbb{R}^{n\times n}$ and $\sigma = [\sigma_{ij}] \in \mathbb{R}^{n\times n}$.

2.3. The financial market

With the notion of cointegration in mind, we consider a financial market in which $n + 1$ assets are traded continuously within the time horizon $[0, T]$. These assets are labeled by S_i for $i = 0, 1, 2 \ldots, n$, with the 0th asset being risk-free. The risk-free asset satisfies the differential equation, $dS_0(t) = (t)S_0(t)dt$, with $S_0(0) = R_0 > 0$, where $r(t)$ is the time deterministic risk-free rate. The risky assets are defined through their log-price processes $X_1(t), \ldots, X_n(t)$, where $X_j(t) = \ln S_j(t)$. The vector of log-prices, $X(t)$, satisfies the SDE

$$dX(t) = [\theta(t) - AX(t)]\,dt + \sigma(t)dW_t, \; t \in [0, T], \quad (3)$$

where $W_t = (W_t^1, \ldots, W_t^n)'$ is a standard $\mathcal{F}_{t\geq 0}$-adapted n-dimensional Wiener process on a fixed filtered complete probability space $(\Omega, \mathcal{F}, \mathcal{P}, \mathcal{F}_{t\geq 0})$, W_t^i and W_t^j are mutually independent for all $i \neq j$, A is an $n \times n$ constant matrix of cointegration coefficients, and $\sigma(t)\sigma(t)' := \Sigma(t)$ is the covariance matrix of assets defined in the Banach space of $\mathbb{R}^{n\times n}$-valued continuous function on $[0, T]$. In line with the literature, we assume that the non-degeneracy condition of $\Sigma(t) \succcurlyeq \delta I_n$ holds for all $t \in [0, T]$ and for some $\delta > 0$. We also assume that $\theta(t)$ and $\sigma(t)$ are measurable and uniformly bounded in $[0, T]$.

2.4. Robust optimal investment

Consider an investor with an initial wealth of Y_0 in the specified financial market with cointegration. The investor seeks an admissible portfolio strategy so that the expected utility of the terminal wealth level is maximized. Let $u_i(t)$ be the amount invested in asset i and $N_i(t)$ be the number of shares of the ith asset in the portfolio of the investor. The wealth of the investor at time t is then defined as $Y(t) = \sum_{i=0}^{n}u_i(t) = \sum_{i=0}^{n}N_i(t)S_i(t)$. Applying Itô's lemma to $Y(t)$ with respect to the cointegrating dynamics (3), the wealth process is given by

$$dY(t) = \left[r(t)Y(t) + u(t)'\mathcal{B}(t)\right]dt + u(t)'\sigma(t)dW_t, \quad (4)$$

with $Y(0) = Y_0$, where

$$\mathcal{B}(t) = \theta - AX(t) + \frac{1}{2}\mathcal{D}(\sigma(t)\sigma(t)')\mathbf{1} - r(t)\mathbf{1}, \quad (5)$$

in which $\mathbf{1}$ is the column vector with all elements being 1, $\mathcal{D}(\sigma(t)\sigma(t)')$ is the diagonal matrix with all diagonal elements equal to those of $\sigma(t)\sigma(t)'$.

The classical utility maximization problem with a fixed investment horizon $T < \infty$ reads $\max_{u(\cdot)}\mathrm{E}[U(Y(T))]$ subject to

[1] Mathematical proofs are collected in the on-line appendix: www.sta.cuhk.edu.hk/hywong/orl2018appendix.pdf.