May 2016

Optimal Pairs Trading Rules

Eric Müller
University of Wisconsin-Milwaukee

OPTIMAL PAIRS TRADING RULES

by

Eric Mueller

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

in MATHEMATICS

at

The University of Wisconsin-Milwaukee

May 2016

ABSTRACT

OPTIMAL PAIRS TRADING RULES

by

Eric Mueller

The University of Wisconsin-Milwaukee, 2016
Under the Supervision of Professor Chao Zhu

This thesis derives an optimal trading rule for a pair of historically correlated stocks. When one stock's price increases and the other one's decreases, a trade of the pair is triggered. The idea is to short the winner and to long the loser with the hope that the prices of the two assets will converge again. In this thesis the spread of the two stocks is governed by a mean-reverting model. The objective is to trade the pair in such a way as to maximize an overall return. The same slippage cost is imposed on every trade. Furthermore, a local-time process to the spread is introduced in order to avoid infinitely large gains.

We use the associated Hamilton-Jacobi-Bellman equations to characterize the value functions which are solved by using the smooth-fit method. It is shown that the solution of the optimal pairs trading problem can be obtained by solving a set of nonlinear equations. Additionally, a set of sufficient conditions is provided in form of a verification theorem. The thesis concludes with a numerical example.

TABLE OF CONTENTS

LIST OF FIGURES

1 Introduction

The research on optimal trading rules may be started from Øksendal (2003) (Example 10.2.2., pp. 219 and Example 10.4.2., pp. 227), where an optimal selling rule for a stock holder was studied. Guo and Zhang (2005) extended the case to a regime switching market, where the stock prices were driven by a geometric Brownian motion combined with a Markov chain. Using a smooth-fit approach they discovered that the optimal trading rules are of a threshold type. Zhang and Zhang (2008) studied the optimal trading rules for both buying and selling in a mean reverting market.

Pairs trading, a convergence trading strategy, involves identifying two stocks whose prices showed similar behavior over a long period of time, i.e., they are historically correlated. If the spread of the two asset prices increases one buys the loser and shorts the winner, betting that history repeats and the prices eventually converge again. This trading strategy was developed in the mid-80's and has been a popular tool used by hedge funds and investment banks since then. For being a successful strategy, it is of great importance to know when to initiate the pairs trade and when to close all positions. The objective of this thesis is to find such rules and establish their optimality. As in Song and Zhang (2013) we consider a mean-reverting model. However, the state process used in this thesis is the difference of the log- prices in contrast to the difference of the real prices used by Song and Zhang. Because of the proportional slippage cost for each transaction in our formulation (see equation (10) for details), and the fact that the difference of the prices of the underlying stocks can be negative, it is possible to have a risk-free positive profit. Therefore in order to work with well-posed problem, we introduce a local time process to the spread, which ensures that the difference of the prices is bounded from below.

The thesis is structured as follows. In Section 2 the pairs trading model is introduced and is followed by the formulation of the optimization problem. In addition, we define the reward function and state important properties of the value function at the end of the section. As in Song and Zhang (2013) and Zhang and Zhang (2008) we follow a dynamic programming approach to solve the optimal stopping time problem. The associated Hamilton-Jacobi-Bellman equations for the value function are established in Section 3, where we also solve them with the help of the smooth-fit method. It is shown that the three threshold levels z_0^*, z_0 and z_1 can be used to construct an optimal trading times. These levels are obtained by solving a set of nonlinear equations. Additionally, we provide sufficient conditions for their optimality in terms of a verification theorem which is proven in Section 4. We conclude the thesis by giving a numerical example in Section 5 which shows the practicability of our computations in the previous sections.

2 Pairs trading model

Let $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\}$ be a complete filtered probability space, in which the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfies the usual condition. Consider two risky asset prices $S_1(t)$ and $S_2(t)$. They are co-integrated in the sense that $X_j(t) := \log(S_j(t))$, $j = 1, 2$ satisfy the stochastic differential equations

$$
\begin{aligned}
dX_1(t) &= [\hat{k}_1 - \theta_1 \hat{Z}(t)]\, dt + \sigma_1\, dW_1(t), \\
dX_2(t) &= [\hat{k}_2 - \theta_2 \hat{Z}(t)]\, dt + \sigma_1\, d\tilde{W}_2(t),
\end{aligned}
\tag{1}
$$

where W_1 and $\tilde{W}_2$ are one-dimensional Brownian motions with correlation coefficient $\rho \in [-1, 1]$, i.e., $\mathbb{E}[dW_1(t)d\tilde{W}_2(t)] = \rho dt$. Let's write $\tilde{W}_2(t) := \rho W_1(t) + \sqrt{1-\rho^2} W_2(t)$, $t \geq 0$, where W_2 is a one-dimensional Brownian motion that is independent of W_1. The stochastic process $\hat{Z}(t)$ satisfies

$$
\hat{Z}(t) = \hat{a} + X_1(t) + cX_2(t), \tag{2}
$$

for some constant c. As a straightforward derivative of the above dynamics, we see

$$
\begin{aligned}
d\hat{Z}(t) &= \Big[(\hat{k}_1 + c\hat{k}_2) - (\theta_1 + c\theta_2)\hat{Z}(t)\Big]\ dt + \sigma_1\, dW_1(t) + c\sigma_2\, d\tilde{W}_2(t) \\
&= \Big[(\hat{k}_1 + c\hat{k}_2) - (\theta_1 + c\theta_2)\hat{Z}(t)\Big]\ dt + [\sigma_1 + c\sigma_2\rho]\, dW_1(t) \\
&+ c\sigma_2\sqrt{1-\rho^2}\, dW_2(t).
\end{aligned}
\tag{3}
$$

Note that if $(\theta_1 + c\theta_2)$ is positive, then $\hat{Z}(t)$ is a mean-reverting process. When $c = -1$, $\hat{a} = 0$, $\hat{Z}(t)$ is the difference between the log-prices of the two risky assets; it is called the spread of the stocks at time t. This case corresponds to a commonly referred pairs trading scenario: Buy the one stock with lower price and sell the one with higher price simultaneously, and close both positions when the lower one gets higher and the higher one gets lower at some time later. In other words: Short the pair when the difference is large and close the position when it is small. Intuitively, this scenario shall work since the difference is following a mean reverting process. We let $c = -1$ for simplicity in this thesis; it is straightforward to extend our results to the case of arbitrary $c \in \mathbb{R}$. Let

$$
\begin{aligned}
k &= \hat{k}_1 - \hat{k}_2, \\
\sigma^2 &= \sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2, \\
\theta &= \theta_1 - \theta_2.
\end{aligned}
\tag{4}
$$

The process $\hat{Z}(t)$ has a long-time average k/θ. For convenience, we will consider the adjusted process $Z(t)$, defined as follows: $Z(t) := \hat{Z}(t) - \frac{k}{\theta}$. Then $Z(t)$ has a long-time mean zero. Moreover, thanks to (2), we have

$$
X_2(t) = \hat{a} - \frac{k}{\theta} + X_1(t) - Z(t) = a + X_1(t) - Z(t), \tag{5}
$$

where $a := \hat{a} - \frac{k}{\theta}$. The corresponding dynamics of X_1, X_2 and Z become

$$
dX_1(t) = (k_1 - \theta_1 Z(t))\, dt + \sigma_1\, dW_1(t), \tag{6}
$$

$$
dX_2(t) = (k_2 - \theta_2 Z(t))\, dt + \rho\sigma_2\, dW_1(t) + \sigma_2\sqrt{1-\rho^2}\, dW_2(t), \tag{7}
$$

$$
dZ(t) = -\theta Z(t)\, dt + (\sigma_1 - \rho\sigma_2)\, dW_1(t) - \sigma_2\sqrt{1-\rho^2}\, dW_2(t), \tag{8}
$$

where $k_i = \hat{k}_i - k\theta_i/\theta$. Assume the initial conditions of (6) and (8) are given by $X_1(0) = x$ and $Z(0) = z$, respectively. Note that thanks to (5), $X_2(0) = a + x - z$ is the initial condition of (7). We will use the above dynamics in the rest of this thesis. We assume $\theta_1 > \theta_2 > 0$. As a result, $\theta = \theta_1 - \theta_2 > 0$. Note the setting implies that the first asset has faster reverting rate hence we shall buy it when the pair (difference) is far away from the long-time average, and short the second one with slower reverting rate. Let us introduce

$$
S(t) = S_1(t) - S_2(t) = e^{X_1(t)} - e^{X_2(t)} = e^{X_1(t)}\left(1 - e^{a - Z(t)}\right), \quad t \geq 0, \tag{9}
$$

where the last equality follows from (5). Note that $S(t)$ can take negative values. Denote

$$\Lambda_0 = (\tau_1, \nu_1, \tau_2, \nu_2, \dots), \quad \Lambda_1 = (\nu_1, \tau_2, \nu_2, \tau_3, \dots),$$

in which τ_i, $i = 1, 2, 3, \dots$ denote the trading times at which we long the pair, i.e. buy S_1 and sell S_2, and ν_i, $i = 1, 2, 3, \dots$ denote the trading times at which we sell the pair. Assume that $\tau_1 < \nu_1 < \tau_2 < \nu_2 < \dots$ are $\mathcal{F}_t$-stopping times. The two sets Λ_0 and Λ_1 represent trading sequences with different first trading types. As required, no short selling of the pair is allowed. It means that we may long the pair or wait for a chance when no positions in hands, and may close all positions or wait when the pair is in hands. Moreover, given initial conditions $X_1(0) = x$, $Z(0) = z$, we define the reward function as follows

$$J_i(z, x, \Lambda_i) = \begin{cases} \mathbb{E}\left[\sum\limits_{n=1}^{\infty} \left(e^{-\alpha\nu_n} S(\nu_n)(1-\delta) - e^{-\alpha\tau_n} S(\tau_n)(1+\delta)\right)\right], & i = 0, \\ \mathbb{E}\left[e^{-\alpha\nu_1} S(\nu_1)(1-\delta) + \sum\limits_{n=2}^{\infty} \left(e^{-\alpha\nu_n} S(\nu_n)(1-\delta) - e^{-\alpha\tau_n} S(\tau_n)(1+\delta)\right)\right], & i = 1, \end{cases} \tag{10}$$

where $\alpha \in (0, 1)$ is the discount factor , and $\delta \in (0, 1)$ is the transaction rate or slippage cost. Without loss of generality, we assume the same rate for buying and selling. For simplicity, the term $\mathbb{E} \sum\limits_{n=1}^{\infty} Y_n$, for an arbitrary sequence of random variables Y_n, will be interpreted as

$$\limsup_{N\to\infty} \mathbb{E} \sum_{n=1}^{\mathbb{N}} Y_n$$

throughout this thesis. The initial net position is represented by i: $i = 0$ means we have no position in hands while $i = 1$ means we have a long position in the pair. As such defined, J_0 is the expected value of cumulative discounted gain (loss) excess transaction cost, given no position in hands initially and J_1 is the expected value, given a current position of the pair in hands. Since $S(t)$ can take negative values there is the possibility to gain an infinitely large profit just by waiting long enough. To exclude this case one could introduce a stop-loss level $\kappa > 0$, as in Song and Zhang (2013), then a selling decision would have to be made before $S(t)$ reaches that level. Another possibility, which we will use in this thesis, is to introduce a local time process to Z so that $Z(t) \geq a$ for all $t \geq 0$ a.s. In this case, Z is a reflected diffusion process with reflection point a. So we modify the SDE (8) by

$$Z(t) = z - \int_0^t \theta Z(s)\, ds + \int_0^t (\sigma_1 - \rho\sigma_2)\, dW_1(s) - \int_0^t \sigma_2\sqrt{1-\rho^2}\, dW_2(s) + L_a(t),$$

where $z \geq a$ and L_a is the local time process of Z at a; that is, for a.e. $\omega \in \Omega$, $L_a(t)$ satisfies

(i) $Z(t) \geq a$ for all $0 \leq t \leq \infty$,

(ii) $L_a(0) = 0$, $L_a(\cdot)$ is non-decreasing and

(iii) $L_a(\cdot)$ is flat off $\{t \geq 0 : Z(t) = a\}$, i.e, $\int_0^\infty \mathbb{1}_{\{Z(s)>a\}}\, dL_a(s) = 0$.

See, for example, Sections 3.6 and 3.7 of Karatzas and Shreve (1991) for details on local time processes. Consequently $S(t) \geq 0$ for all $t \geq 0$. We use the term "buy the pair" to denote the action of longing one share of S_1 and shorting one share of S_2 simultaneously. Similarly "sell the pair" means to sell one share of S_1 and buy one share of S_2 simultaneously. For $i = 0, 1$ let $V_i(z, x)$ denote the value functions with the initial state $Z(0) = z > a$, $X_1(0) = x$ and the initial net positions of the pair $i = 0, 1$. That is,

$$V_i(z,x) = \sup_{\Lambda_i} J_i(z, x, \Lambda_i). \tag{11}$$

We will solve the optimization problem (11) to find optimal pairs trading rules Λ_i^* for $i = 0, 1$.

2.1 Properties of the value function

First we notice that $\Lambda_0 = (\tau_1, \nu_1, \tau_2, \nu_2, \dots)$ can be interpreted as a combination of a buy at τ_1 followed by the sequence $\Lambda_1 = (\nu_1, \tau_2, \nu_2, \dots)$ starting with a sell. Therefore

$$\begin{aligned}
V_0(z,x) &\geq J_0(z,x,\Lambda_0) \\
&= \mathbb{E}[e^{-\alpha\nu_1} S(\nu_1)(1-\delta) + \sum_{n=2}^{\infty} \big(e^{-\alpha\nu_n} S(\nu_n)(1-\delta) - e^{-\alpha\tau_n} S(\tau_n)(1+\delta)\big) \\
&\quad - e^{-\alpha\tau_1} S(\tau_1)(1+\delta)] \\
&= J_1(Z(\tau_1), X_1(\tau_1), \Lambda_1) - \mathbb{E}e^{-\alpha\tau_1} S(\tau_1)(1+\delta).
\end{aligned}$$

Now by setting $\tau_1 = 0$ and taking the supremum over all Λ_1 we obtain

$$V_0(z,x) \geq V_1(z,x) - e^x(1-e^{a-z})(1+\delta). \tag{12}$$

Similarly, one can regard $\Lambda_1 = (\nu_1, \tau_2, \nu_2, \dots)$ as a combination of a sell at ν_1 followed by a sequence starting with a buy $\Lambda_0 = (\tau_2, \nu_2, \dots)$. Hence

$$\begin{aligned}
V_1(z,x) &\geq J_1(z,x,\Lambda_1) \\
&= \mathbb{E}[e^{-\alpha\nu_1} S(\nu_1)(1-\delta) + \sum_{n=2}^{\infty} \big(e^{-\alpha\nu_n} S(\nu_n)(1-\delta) - e^{-\alpha\tau_n} S(\tau_n)(1+\delta)\big)] \\
&= J_0(Z(\nu_1), X_1(\nu_1), \Lambda_0) + \mathbb{E}e^{-\alpha\nu_1} S(\nu_1)(1-\delta).
\end{aligned}$$

Let $\nu_1 = 0$ and take the supremum over all Λ_0 to get

$$V_1(z,x) \geq V_0(z,x) + e^x(1-e^{a-z})(1-\delta). \tag{13}$$

In the next lemma we will establish bounds for $V_i(z,x)$.

Lemma 1

There exists a constant K_0 such that the following inequalities hold for $z \geq a$ and $x \in \mathbb{R}$

$$0 \leq V_0(z,x) \leq K_0, \tag{14}$$

$$e^x(1-e^{a-z})(1-\delta) \leq V_1(z,x) \leq K_0 + e^x(1-e^{a-z})(1+\delta). \tag{15}$$

Proof. The lower bounds follow from the definitions of $V_i(z,x)$. For the upper bounds consider the following process

$$Y(t) = \begin{pmatrix} X_1(t) \\ Z(t) \end{pmatrix}.$$

Then

$$dY(t) = \begin{pmatrix} k_1 - \theta_1 Z(t) \\ -\theta Z(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ \sigma_1 - \rho\sigma_2 & -\sigma_2\sqrt{1-\rho^2} \end{pmatrix} d \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dL_a(t).$$

Then the generator $\mathcal{A}$ of Y is given by

$$\mathcal{A}f = (k_1 - \theta_1 z)\,\frac{\partial f}{\partial x} - \theta z\,\frac{\partial f}{\partial z} + \frac{1}{2}\sigma_1^2\,\frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\left[(\sigma_1 - \rho\sigma_2)^2 + \sigma_2^2(1-\rho^2)\right]\frac{\partial^2 f}{\partial z^2}$$
$$+ \sigma_1(\sigma_1 - \rho\sigma_2)\,\frac{\partial^2 f}{\partial x \partial z}.$$

Since $\sigma_1 \frac{\partial f}{\partial x}$, $(\sigma_1 - \rho\sigma_2)\frac{\partial f}{\partial z}$ and $-\sigma_2\sqrt{1-\rho^2}\frac{\partial f}{\partial z}$ are all continuous for $f(t,z,x) = e^{-\alpha t}e^x(1-e^{a-z})$, they are bounded on $[0,t]$. Therefore

$$0 = \mathbb{E}\int_0^t \sigma_1 e^{-\alpha s} e^{X_1(s)}(1-e^{a-Z(s)})dW_1(s)$$
$$= \mathbb{E}\int_0^t (\sigma_1 - \rho\sigma_2)e^{-\alpha s}e^{X_1(s)}e^{a-Z(s)}dW_1(s)$$
$$= \mathbb{E}\int_0^t -\sigma_2\sqrt{1-\rho^2}e^{-\alpha s}e^{X_1(s)}e^{a-Z(s)}dW_2(s).$$

Then according to Dynkin's formula, we have

$$\mathbb{E}e^{-\alpha\nu_n}S(\nu_n) - \mathbb{E}e^{-\alpha\tau_n}S(\tau_n) = \mathbb{E}\int_{\tau_n}^{\nu_n}(\mathcal{A}-\alpha)e^{-\alpha t}e^{X_1(t)}(1-e^{a-Z(t)})dt,$$

where

$$\begin{aligned}
&(\mathcal{A}-\alpha)e^{-\alpha t}e^{x}(1-e^{a-z})\\
&= -\alpha e^{-\alpha t}e^{x}(1-e^{a-z}) + (k_1-\theta_1 z)e^{-\alpha t}e^{x}(1-e^{a-z})\\
&- \theta z e^{-\alpha t}e^{x}e^{a-z} + \frac{1}{2}\sigma_1^2 e^{-\alpha t}e^{x}(1-e^{a-z}) - \frac{1}{2}\left[(\sigma_1-\rho\sigma_2)^2+\sigma_2^2(1-\rho^2)\right]e^{-\alpha t}e^{x}e^{a-z}\\
&+ \sigma_1(\sigma_1-\rho\sigma_2)e^{-\alpha t}e^{x}e^{a-z}\\
&= e^{-\alpha t}\Big[e^{x}(1-e^{a-z})(-\alpha+k_1+\frac{1}{2}\sigma_1^2-\theta_1 z)\\
&+ e^{x}e^{a-z}(-\theta z-\frac{1}{2}\sigma_1^2+\sigma_1\sigma_2\rho-\frac{1}{2}\sigma_2^2\rho^2-\frac{1}{2}\sigma_2^2+\frac{1}{2}\sigma_2^2\rho^2+\sigma_1^2-\sigma_1\sigma_2\rho)\Big]\\
&= e^{-\alpha t}\Big[e^{x}(1-e^{a-z})(k_1-\alpha+\frac{1}{2}\sigma_1^2-\theta_1 z)+e^{x}e^{a-z}(\frac{1}{2}\sigma_1^2-\frac{1}{2}\sigma_2^2-\theta z)\Big].
\end{aligned}$$

For $V_0(z,x)$ to have an upper bound we need $(\mathcal{A}-\alpha)e^{-\alpha t}e^{x}(1-e^{a-z})$ to be bounded from above. At first we consider the limits when $(x,z)\to(\infty,\infty)$ and $(x,z)\to(-\infty,\infty)$. Since $z\geq a$, we don't have to examine the cases when $z\to-\infty$.

- $(x,z)\to(\infty,\infty)$: Since e^x and $(1-e^{a-z})$ are positive and $(k_1-\alpha+\frac{1}{2}\sigma_1^2-\theta_1 z)$ becomes negative for z large enough the limit of $e^x(1-e^{a-z})(k_1-\alpha+\frac{1}{2}\sigma_1^2-\theta_1 z)$ is negative. Similar for $e^xe^{a-z}(\frac{1}{2}\sigma_1^2-\frac{1}{2}\sigma_2^2-\theta z)$.

- $(x,z)\to(-\infty,\infty)$: e^x and e^{a-z} converge to 0 and $(1-e^{a-z})$ to 1 but always stays positive. However, as before $(k_1-\alpha+\frac{1}{2}\sigma_1^2-\theta_1 z)$ and $(\frac{1}{2}\sigma_1^2-\frac{1}{2}\sigma_2^2-\theta z)$ become negative, which means that the limits of both $e^x(1-e^{a-z})(k_1-\alpha+\frac{1}{2}\sigma_1^2-\theta_1 z)$ and $e^xe^{a-z}(\frac{1}{2}\sigma_1^2-\frac{1}{2}\sigma_2^2-\theta z)$ are negative.

- Because of the connection between $Z(t)$ and $X_1(t)$, z converges to infinity when x does and vice versa. So all in all $(\mathcal{A}-\alpha)e^{-\alpha t}e^{x}(1-e^{a-z})$ cannot get infinitely large and therefore is bounded from above.

Let C be an upper bound. Then

$$\mathbb{E}e^{-\alpha\nu_n}S(\nu_n) - \mathbb{E}e^{-\alpha\tau_n}S(\tau_n) \leq C \cdot \mathbb{E}\int_{\tau_n}^{\nu_n} e^{-\alpha t}dt.$$

With the definition of $J_0(z, x, \Lambda_0)$ it follows

$$\begin{aligned}
J_0(z,x,\Lambda_0) =& \mathbb{E}\sum_{n=1}^{\infty}\left(e^{-\alpha\nu_n}S(\nu_n)(1-\delta) - e^{-\alpha\tau_n}S(\tau_n)(1+\delta)\right) \\
=& \mathbb{E}\sum_{n=1}^{\infty}\left(e^{-\alpha\nu_n}S(\nu_n) - e^{-\alpha\tau_n}S(\tau_n)\right) - \underbrace{\delta\left(e^{-\alpha\nu_n}S(\nu_n) + e^{-\alpha\tau_n}S(\tau_n)\right)}_{\geq 0} \\
\leq& \sum_{n=1}^{\infty}\left(\mathbb{E}e^{-\alpha\nu_n}S(\nu_n) - \mathbb{E}e^{-\alpha\tau_n}S(\tau_n)\right) \\
\leq& \sum_{n=1}^{\infty} C \cdot \mathbb{E}\int_{\tau_n}^{\nu_n} e^{-\alpha t}dt \leq C\int_{0}^{\infty} e^{-\alpha t}dt = \frac{C}{\alpha} := K_0.
\end{aligned}$$

Thus, $0 \leq V_0(z,x) \leq K_0$. Since $V_0(z,x) \geq V_1(z,x) - e^x(1-e^{a-z})(1+\delta)$ we have $e^x(1-e^{a-z})(1-\delta) \leq V_1(z,x) \leq K_0 + e^x(1-e^{a-z})(1-\delta)$. This completes the proof. $\square$

3 HJB equation and its solution

Following the dynamic programming method to solve the considered stochastic optimization problem, the associated Hamilton-Jacobi-Bellman (HJB) equations are formally given by

$$
\begin{cases}
\min\{(\alpha - \mathcal{A})v_0(z,x),\ v_0(z,x) - v_1(z,x) + e^x(1-e^{a-z})(1+\delta)\} = 0, \\
\min\{(\alpha - \mathcal{A})v_1(z,x),\ v_1(z,x) - v_0(z,x) - e^x(1-e^{a-z})(1-\delta)\} = 0, \\
\frac{\partial v_0}{\partial z}(a,x) \le 0, \quad \frac{\partial v_1}{\partial z}(a,x) = 0.
\end{cases}
$$

If $i = 0$, i.e., no position at hand, intuitively, one should buy when the spread is small (say equal or less than $z_0 > a$). Then the continuation region, on which $(\alpha - \mathcal{A})v_0(z,x) = 0$, should include (z_0, ∞). Furthermore, because of the mean reverting character of $Z(t)$ it makes sense not to buy when the spread is close to a. We want to be sure that the prices really diverge and not immediately return to the equilibrium level. Therefore the continuation region should additionally include the interval (a, z_0^*) for $a < z_0^* < z_0$. Then the action region is given by (z_0^*, z_0) on which we have $v_0(z,x) = v_1(z,x) - e^x(1-e^{a-z})(1+\delta)$. On the other hand, if $i = 1$, i.e., a position at hand, one should sell if the spread is large (say equal or greater than $z_1 > z_0$). Then the continuation region is given by (a, z_1), on which one should have $(\alpha - \mathcal{A})v_1(z,x) = 0$. Consequently, for $z > z_1$ (the action region) we have $v_1(z,x) = v_0(z,x) + e^x(1-e^{a-z})(1-\delta)$. These regions are illustrated in Fig. 1.

First of all we try to solve the equations $(\alpha - \mathcal{A})v_i(z,x) = 0$, $i = 0, 1$. Suppose a

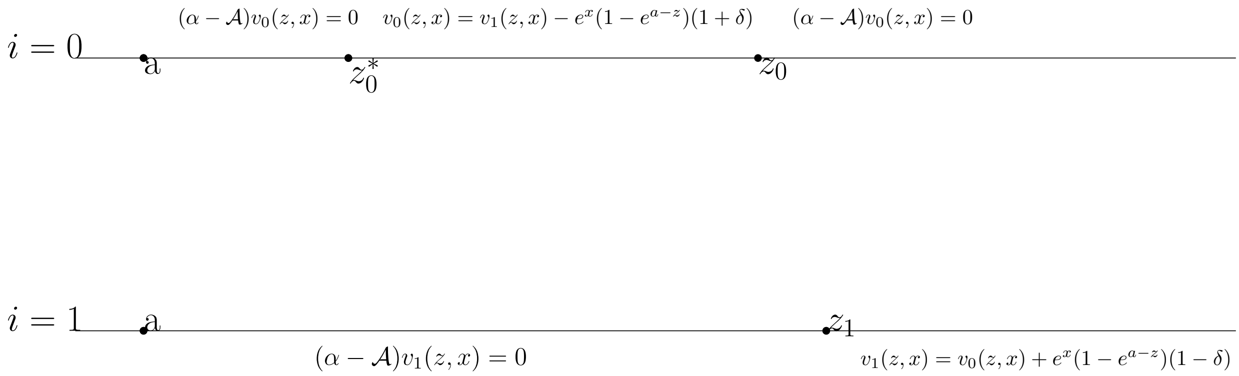

Figure 1: Continuation and action regions

possible solution has the form $u(z,x) = e^x g(z)$. Thus

$$
\begin{aligned}
&(\alpha - \mathcal{A})u(z,x) = 0 \\
\Leftrightarrow\ & \alpha u(z,x) - (k_1 - \theta_1 z)u_x(z,x) + \theta z u_z(z,x) - \frac{1}{2}\sigma_1^2 u_{xx}(z,x) \\
&- \frac{1}{2}(\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2)u_{zz}(z,x) - \sigma_1(\sigma_1 - \sigma_2\rho)u_{zx}(z,x) = 0 \\
\Leftrightarrow\ & \alpha e^x g(z) - (k_1 - \theta_1 z)e^x g(z) + \theta z e^x g'(z) - \frac{1}{2}\sigma_1^2 e^x g(z) \\
&- \frac{1}{2}(\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2)e^x g''(z) - \sigma_1(\sigma_1 - \sigma_2\rho)e^x g(z) = 0 \\
\Leftrightarrow\ & e^x\big(g(z)[\alpha - \frac{1}{2}\sigma_1^2 - k_1 + \theta_1 z] + g'(z)[-\sigma_1^2 + \sigma_1\sigma_2\rho + \theta z] \\
&- g''(z)\frac{1}{2}[\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2]\big) = 0 \\
\Leftrightarrow\ & g''(z) + \frac{2(\sigma_1^2 - \sigma_1\sigma_2\rho - \theta z)}{(\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2)}g'(z) + \frac{2(\frac{1}{2}\sigma_1^2 + k_1 - \alpha - \theta_1 z)}{(\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2)}g(z) = 0.
\end{aligned}
\tag{16}
$$

Next define

$$
\begin{aligned}
A &:= \frac{2(\sigma_1^2 - \sigma_1\sigma_2\rho)}{\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2}, \quad B := \frac{-2\theta}{\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2} < 0, \\
C &:= \frac{2(\frac{1}{2}\sigma_1^2 + k_1 - \alpha)}{\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2}, \quad D := \frac{-2\theta_1}{\sigma_1^2 - 2\sigma_1\sigma_2\rho + \sigma_2^2} < 0.
\end{aligned}
$$

Then equation (16) becomes

$$
g''(z) + (A + Bz)g'(z) + (C + Dz)g(z) = 0
$$

and the solution is given by

$$g(z) = C_1 e^{-\frac{D}{B}z} KummerM\left(\frac{1}{2}\frac{B^2C + D^2 - ABD}{B^3}, \frac{1}{2}, -\frac{1}{2}\frac{B^2z + AB - 2D^2}{B^3}\right)$$
$$+ C_2 e^{-\frac{D}{B}z} KummerU\left(\frac{1}{2}\frac{B^2C + D^2 - ABD}{B^3}, \frac{1}{2}, -\frac{1}{2}\frac{B^2z + AB - 2D^2}{B^3}\right).$$

For some constants C_1 and C_2. Note that $D/B = \theta_1/\theta$. Hence

$$g(z) = C_1 e^{-\frac{\theta_1}{\theta}z} KummerM\left(\frac{1}{2}\frac{B^2C + D^2 - ABD}{B^3}, \frac{1}{2}, -\frac{1}{2}\frac{B^2z + AB - 2D^2}{B^3}\right)$$
$$+ C_2 e^{-\frac{\theta_1}{\theta}z} KummerU\left(\frac{1}{2}\frac{B^2C + D^2 - ABD}{B^3}, \frac{1}{2}, -\frac{1}{2}\frac{B^2z + AB - 2D^2}{B^3}\right).$$

For simplicity let the following

$$\phi_1(z) := e^{-\frac{\theta_1}{\theta}z} KummerM\left(\frac{1}{2}\frac{B^2C + D^2 - ABD}{B^3}, \frac{1}{2}, -\frac{1}{2}\frac{B^2z + AB - 2D^2}{B^3}\right),$$
$$\phi_2(z) := e^{-\frac{\theta_1}{\theta}z} KummerU\left(\frac{1}{2}\frac{B^2C + D^2 - ABD}{B^3}, \frac{1}{2}, -\frac{1}{2}\frac{B^2z + AB - 2D^2}{B^3}\right).$$

With respect to the continuation regions and our value functions this means there exist constants A_1, A_2, B_1, B_2 and C_1, C_2 such that $v_0(z, x) = e^x(A_1\phi_1(z) + A_2\phi_2(z))$ on (a, z_0^*) and $v_0(z, x) = e^x(C_1\phi_1(z) + C_2\phi_2(z))$ on (z_0, ∞) and $v_1(z, x) = e^x(B_1\phi_1(z) + B_2\phi_2(z))$ on (a, z_1). Consider first the interval (z_1, ∞), since according to Lemma 1 $v_0(z, x)$ is bounded from above and $\phi_1(z) \xrightarrow[z\to\infty]{} \infty$ this implies $C_1 = 0$ and $v_0(z, x) = e^x C_2\phi_2(z)$. So in total

$$v_0(z, x) = \begin{cases} e^x C_2\phi_2(z) & \text{on } (z_0, \infty), \\ e^x(A_1\phi_1(z) + A_2\phi_2(z)) & \text{on } (a, z_0^*), \end{cases}$$

and

$$v_1(z, x) = e^x(B_1\phi_1(z) + B_2\phi_2(z)) \text{ on } (a, z_1).$$

It is easy to see that both v_0 and v_1 are twice continuously differentiable on their continuation regions. We want to apply the smooth-fit method which requires the solutions to be continuously differentiable

- at z_0^*:

$$
\begin{aligned}
v_0(z_0^*, x) &= v_1(z_0^*, x) - e^x(1 - e^{a - z_0^*})(1+\delta), \\
\left.\frac{\partial v_0(z,x)}{\partial z}\right|_{z_0^*} &= \left.\frac{\partial v_1(z,x)}{\partial z}\right|_{z_0^*} - e^x e^{a - z_0^*}(1+\delta),
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
A_1\phi_1(z_0^*) + A_2\phi_2(z_0^*) &= B_1\phi_1(z_0^*) + B_2\phi_2(z_0^*) - (1 - e^{a - z_0^*})(1+\delta), \\
A_1\phi_1'(z_0^*) + A_2\phi_2'(z_0^*) &= B_1\phi_1'(z_0^*) + B_2\phi_2'(z_0^*) - e^{a - z_0^*}(1+\delta).
\end{aligned} \tag{17}
$$

- at z_0:

$$
\begin{aligned}
v_0(z_0, x) &= v_1(z_0, x) - e^x(1 - e^{a - z_0})(1+\delta), \\
\left.\frac{\partial v_0(z,x)}{\partial z}\right|_{z_0} &= \left.\frac{\partial v_1(z,x)}{\partial z}\right|_{z_0} - e^x e^{a - z_0}(1+\delta),
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
C_2\phi_2(z_0) &= B_1\phi_1(z_0) + B_2\phi_2(z_0) - (1 - e^{a - z_0})(1+\delta), \\
C_2\phi_2'(z_0) &= B_1\phi_1'(z_0) + B_2\phi_2'(z_0) - e^{a - z_0}(1+\delta).
\end{aligned} \tag{18}
$$

- at z_1:

$$
\begin{aligned}
v_1(z_1, x) &= v_0(z_1, x) + e^x(1 - e^{a - z_0})(1-\delta), \\
\left.\frac{\partial v_1(z,x)}{\partial z}\right|_{z_1} &= \left.\frac{\partial v_0(z,x)}{\partial z}\right|_{z_1} + e^x e^{a - z_0}(1-\delta),
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
&B_1\phi_1(z_1) + B_2\phi_2(z_1) = C_2\phi_2(z_1) + (1 - e^{a-z_1})(1-\delta), \\
&B_1\phi_1'(z_1) + B_2\phi_2'(z_1) = C_2\phi_2'(z_1) + e^{a-z_1}(1-\delta).
\end{aligned} \tag{19}
$$

- Additionally, we need v_0 and v_1 to satisfy the following at a:

$$
\left.\frac{\partial v_0(z,x)}{\partial z}\right|_a = 0,
$$
$$
\left.\frac{\partial v_1(z,x)}{\partial z}\right|_a = 0,
$$

which is equivalent to

$$
\begin{aligned}
&A_1\phi_1'(a) + A_2\phi_2'(a) = 0, \\
&B_1\phi_1'(a) + B_2\phi_2'(a) = 0.
\end{aligned} \tag{20}
$$

For simplicity define

$$
\Phi(z) = \begin{pmatrix} \phi_1(z) & \phi_2(z) \\ \phi_1'(z) & \phi_2'(z) \end{pmatrix}
$$

and assume that it is invertible. Then we can now rewrite equations (17)-(20) in terms of $\Phi(z)$:

$$
\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} - (1+\delta)\Phi^{-1}(z_0^*) \begin{pmatrix} 1 - e^{a-z_0^*} \\ e^{a-z_0^*} \end{pmatrix}, \tag{21}
$$

$$
C_2\Phi^{-1}(z_0) \begin{pmatrix} \phi_2(z_0) \\ \phi_2'(z_0) \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} - (1+\delta)\Phi^{-1}(z_0) \begin{pmatrix} 1 - e^{a-z_0} \\ e^{a-z_0} \end{pmatrix}, \tag{22}
$$

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = C_2 \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix} + (1-\delta)\Phi^{-1}(z_1) \begin{pmatrix} 1 - e^{a-z_1} \\ e^{a-z_1} \end{pmatrix}, \tag{23}$$

$$\begin{aligned} &(\phi_1'(a), \phi_2'(a)) \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = 0, \\ &(\phi_1'(a), \phi_2'(a)) \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = 0. \end{aligned} \tag{24}$$

Now multiply both sides of equation (21) by $(\phi_1'(a), \phi_2'(a))$ from the left and use equation (24) to obtain

$$(\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_0^*) \begin{pmatrix} 1 - e^{a-z_0^*} \\ e^{a-z_0^*} \end{pmatrix} = 0. \tag{25}$$

Next combine equations (22) and (23)

$$\begin{aligned} &C_2 \left[\Phi^{-1}(z_0) \begin{pmatrix} \phi_2(z_0) \\ \phi_2'(z_0) \end{pmatrix} - \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix} \right] \\ &= (1-\delta)\Phi^{-1}(z_1) \begin{pmatrix} 1 - e^{a-z_1} \\ e^{a-z_1} \end{pmatrix} - (1+\delta)\Phi^{-1}(z_0) \begin{pmatrix} 1 - e^{a-z_0} \\ e^{a-z_0} \end{pmatrix}. \end{aligned} \tag{26}$$

By multiplying both sides of equation (23) by $(\phi_1'(a), \phi_2'(a))$ from the left and again by using equation (24) we obtain

$$0 = C_2\, (\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix} + (1-\delta)\, (\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} 1 - e^{a-z_1} \\ e^{a-z_1} \end{pmatrix}, \tag{27}$$

which leads to, provided that $(\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix} \neq 0$,

$$C_2 = \frac{-(1-\delta)\,(\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} 1 - e^{a-z_1} \\ e^{a-z_1} \end{pmatrix}}{(\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix}}. \tag{28}$$

Finally, plug this into equation (26) to get

$$\begin{aligned} &\frac{-(1-\delta)\,(\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} 1 - e^{a-z_1} \\ e^{a-z_1} \end{pmatrix}}{(\phi_1'(a), \phi_2'(a))\, \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix}} \\ &\times \left[\Phi^{-1}(z_0) \begin{pmatrix} \phi_2(z_0) \\ \phi_2'(z_0) \end{pmatrix} - \Phi^{-1}(z_1) \begin{pmatrix} \phi_2(z_1) \\ \phi_2'(z_1) \end{pmatrix} \right] \\ = \; &(1-\delta)\Phi^{-1}(z_1) \begin{pmatrix} 1 - e^{a-z_1} \\ e^{a-z_1} \end{pmatrix} - (1+\delta)\Phi^{-1}(z_0) \begin{pmatrix} 1 - e^{a-z_0} \\ e^{a-z_0} \end{pmatrix}. \end{aligned} \tag{29}$$

First, one can obtain the triple (z_0^*, z_0, z_1) by solving the equations (25) and (29). In order to get the constants C_2, A_1, A_2 and B_1, B_2 one has to solve first equation (28) and then (21) and (23).

Furthermore, we need additional requirements for $v_i(z,x)$. The value functions have to satisfy the following conditions for being solutions to the HJB-equations:

$$\begin{aligned} &(\alpha - \mathcal{A})v_0(z,x) \geq 0, \\ &(\alpha - \mathcal{A})v_1(z,x) \geq 0, \\ &v_0(z,x) \geq v_1(z,x) - e^x(1 - e^{a-z})(1+\delta), \\ &v_1(z,x) \geq v_0(z,x) + e^x(1 - e^{a-z})(1-\delta). \end{aligned}$$

Let us examine these inequalities on the intervals (a, z_0^*), (z_0^*, z_0), (z_0, z_1) and (z_1, ∞). On (a, z_0^*), the first two inequalities become equalities. Therefore, only the last to inequalities have to hold which is equivalent to

$$e^x(1-e^{a-z})(1-\delta) \leq v_1(z,x) - v_0(z,x) \leq e^x(1-e^{a-z})(1+\delta). \tag{30}$$

On (z_0^*, z_0), we have $v_0(z,x) = v_1(z,x) - e^x(1-e^{a-z})(1+\delta)$ which implies $v_1(z,x) \geq v_0(z,x) + e^x(1-e^{a-z})(1-\delta)$. Hence we only need $(\alpha - \mathcal{A})v_0(z,x) \geq 0$ to hold since $(\alpha - \mathcal{A})v_1(z,x) = 0$ on (a, z_1). We use $v_0(z,x) = v_1(z,x) - e^x(1-e^{a-z})(1+\delta)$ and $(\alpha - \mathcal{A})v_1(z,x) = 0$ to obtain

$$\begin{aligned}
&(\alpha - \mathcal{A})v_0(z,x) = (\alpha - \mathcal{A})(v_1(z,x) - e^x(1-e^{a-z})(1+\delta)) \\
&= (\alpha - \mathcal{A})v_1(z,x) - (\alpha - \mathcal{A})e^x(1-e^{a-z})(1+\delta) \\
&= \left[e^x(1-e^{a-z})(k_1 - \alpha + \frac{1}{2}\sigma_1^2 - \theta_1 z) + e^x e^{a-z}(\frac{1}{2}\sigma_1^2 - \frac{1}{2}\sigma_2^2 - \theta z)\right](1+\delta) \geq 0,
\end{aligned}$$

which is equivalent to

$$(1-e^{a-z})(k_1 - \alpha + \frac{1}{2}\sigma_1^2 - \theta_1 z) + e^{a-z}(\frac{1}{2}\sigma_1^2 - \frac{1}{2}\sigma_2^2 - \theta z) \geq 0. \tag{31}$$

On (z_0, z_1) the first two inequalities are already fulfilled and similar to (a, z_0^*) we need $v_i(z,x)$ to satisfy

$$e^x(1-e^{a-z})(1-\delta) \leq v_1(z,x) - v_0(z,x) \leq e^x(1-e^{a-z})(1+\delta). \tag{32}$$

Finally, On (z_1, ∞), we have $v_1(z,x) = v_0(z,x) + e^x(1-e^{a-z})(1-\delta)$ which implies $v_0(z,x) \geq v_1(z,x) - e^x(1-e^{a-z})(1+\delta)$. Hence we only need $(\alpha - \mathcal{A})v_1(z,x) \geq 0$ to hold since $(\alpha - \mathcal{A})v_0(z,x) = 0$ on (z_0, ∞). We use $v_1(z,x) = v_0(z,x) + e^x(1-e^{a-z})(1-\delta)$ and $(\alpha - \mathcal{A})v_0(z,x) = 0$ to obtain

$$\begin{aligned}
&(\alpha - \mathcal{A})v_1(z,x) = (\alpha - \mathcal{A})(v_0(z,x) + e^x(1-e^{a-z})(1-\delta)) \\
&= (\alpha - \mathcal{A})v_0(z,x) + (\alpha - \mathcal{A})e^x(1-e^{a-z})(1-\delta) \\
&= \left[e^x(1-e^{a-z})(-k_1 + \alpha - \frac{1}{2}\sigma_1^2 + \theta_1 z) + e^x e^{a-z}(-\frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2 + \theta z)\right](1-\delta) \geq 0,
\end{aligned}$$

which is equivalent to

$$(1-e^{a-z})(k_1-\alpha+\frac{1}{2}\sigma_1^2-\theta_1 z)+e^{a-z}(\frac{1}{2}\sigma_1^2-\frac{1}{2}\sigma_2^2-\theta z)\le 0. \tag{33}$$

Note that the inequalities in (30) and (32) are equivalent to the following

$$\begin{aligned}&\left|\frac{(B_1-A_1)\phi_1(z)+(B_2-A_2)\phi_2(z)-(1-e^{a-z})}{1-e^{a-z}}\right|\le\delta \quad \text{on } (a,z_0^*),\\ &\left|\frac{B_1\phi_1(z)+(B_2-C_2)\phi_2(z)-(1-e^{a-z})}{1-e^{a-z}}\right|\le\delta \quad \text{on } (z_0,z_1).\end{aligned} \tag{34}$$

In the following section we show that the triple (z_0^*, z_0, z_1) satisfying the conditions above can be used to construct the optimal trading rules.

4 Verification theorem

In this section we show that the triple (z_0^*, z_0, z_1) satisfying the conditions in Section 3 can be used to construct optimal trading rules. In addition, we show that the functions $v_i(z,x), i = 0,1$ given in Section 3 are equal to the value functions $V_i(z,x), i = 0,1$ defined in (11).

Theorem 2 (Verification Theorem)

Let (z_0^*, z_0, z_1) be a solution to (25) and (29) such that

$$\begin{aligned} &(1-e^{a-z})(k_1 - \alpha + \frac{1}{2}\sigma_1^2 - \theta_1 z) + e^{a-z}(\frac{1}{2}\sigma_1^2 - \frac{1}{2}\sigma_2^2 - \theta z) \geq 0 \quad \forall z \in (z_0^*, z_0), \\ &(1-e^{a-z})(k_1 - \alpha + \frac{1}{2}\sigma_1^2 - \theta_1 z) + e^{a-z}(\frac{1}{2}\sigma_1^2 - \frac{1}{2}\sigma_2^2 - \theta z) \leq 0 \quad \forall z \in (z_1, \infty). \end{aligned} \tag{35}$$

Furthermore, let A_1, A_2, B_1, B_2 and C_2 be constants given by (28), (21) and (23) satisfying (34). Let

$$v_0(z,x) = \begin{cases} e^x(A_1\phi_1(z) + A_2\phi_2(z)) & \textit{on } [a, z_0^*), \\ e^x(B_1\phi_1(z) + B_2\phi_2(z) - (1-e^{a-z})(1+\delta)) & \textit{on } [z_0^*, z_0), \\ e^x C_2\phi_2(z) & \textit{on } [z_0, \infty), \end{cases}$$

$$v_1(z,x) = \begin{cases} e^x(B_1\phi_1(z) + B_2\phi_2(z)) & \textit{on } [a, z_1), \\ e^x(C_2\phi_2(z) + (1-e^{a-z})(1-\delta)) & \textit{on } [z_1, \infty). \end{cases}$$

Moreover, assume $v_0(z,x) \geq 0$. Then, $v_i(z,x) = V_i(z,x)$, $i = 0,1$. Additionally, if $i = 0$, let

$$\Lambda_0^* = (\tau_1^*, \nu_1^*, \tau_2^*, \nu_2^*, \dots),$$

where the stopping times τ_i^* and ν_i^*, $i = 1, 2, \ldots$ are defined in the following way for $n \geq 1$

$$\tau_i^* = \begin{cases} \inf\{t \geq 0 : z_0^* \leq Z(t) \leq z_0\}, & \text{if } i = 1, \\ \inf\{t > \nu_n^* : z_0^* \leq Z(t) \leq z_0\}, & \text{if } i = n+1, \end{cases}$$

$$\nu_n^* = \inf\{t > \tau_n^* : z_1 < Z(t)\}.$$

Likewise, if $i = 1$, let

$$\Lambda_1^* = (\nu_1^*, \tau_2^*, \nu_2^*, \tau_3^*, \ldots),$$

where for $n \geq 2$

$$\nu_i^* = \begin{cases} \inf\{t \geq 0 : z_1 < Z(t)\}, & \text{if } i = 1, \\ \inf\{t > \tau_n^* : z_1 < Z(t)\}, & \text{if } i = n, \end{cases}$$

$$\tau_n^* = \inf\{t > \nu_{n-1}^* : z_0^* \leq Z(t) \leq z_0\}.$$

Then Λ_0^* and Λ_1^* are optimal.

Proof. The proof of the theorem consists of two parts. First, we show that $v_i(z, x) \geq J_i(z, x, \Lambda_i)$, $i = 0, 1$ for all $x \in \mathbb{R}$ and $z \geq a$. Subsequently, we prove $v_i(z, x) = J_i(z, x, \Lambda_i^*)$ which implies $v_i(z, x) = V_i(z, x)$ and the optimality of Λ_i^*.
At first denote $I_0 = (a, z_0^*) \cup (z_0^*, z_0) \cup (z_0, \infty)$ and $I_1 = (a, z_1) \cup (z_1, \infty)$. It is easy to see that $v_0(z, x) \in C^2(I_0 \times \mathbb{R})$ and $v_1(z, x) \in C^2(I_1 \times \mathbb{R})$. Additionally, both are in $C^1([a, \infty])$. Furthermore, v_1 and v_2 satisfy $(\alpha - \mathcal{A})v_i(z, x) \geq 0$ on I_0 and I_1, respectively.With these inequalities and Dynkin's formula, we have for any stopping times $0 \leq \sigma_1 \leq \sigma_2$

$$\begin{aligned} &\mathbb{E}e^{-\alpha\sigma_2} v_i(Z(\sigma_2), X_1(\sigma_2)) - \mathbb{E}e^{-\alpha\sigma_1} v_i(Z(\sigma_1), X_1(\sigma_1)) \\ =&\mathbb{E}\int_{\sigma_1}^{\sigma_2} e^{-\alpha t} \underbrace{(-\alpha + \mathcal{A})v_i(Z(t), X(t))}_{\leq 0} dt \leq 0. \end{aligned}$$

Note that since $v_i(z, x) \in C^2$ and $\frac{\partial v_i(a,x)}{\partial z} = 0$, Dynkin's formula is applicable in this situation. Hence,

$$\mathbb{E}e^{-\alpha\sigma_2}v_i(Z(\sigma_2),X_1(\sigma_2)) \leq \mathbb{E}e^{-\alpha\sigma_1}v_i(Z(\sigma_1),X_1(\sigma_1)). \tag{36}$$

Since $\tau_1 \geq 0$, this implies

$$\begin{aligned}
v_0(z,x) &= \mathbb{E}e^{-\alpha\cdot 0}v_0(Z(0),X_1(0)) \geq \mathbb{E}e^{-\alpha\tau_1}v_0(Z(\tau_1),X_1(\tau_1)) \\
&\geq \mathbb{E}e^{-\alpha\tau_1}\left[v_1(Z(\tau_1),X_1(\tau_1)) - S(\tau_1)(1+\delta)\right] \\
&= \mathbb{E}e^{-\alpha\tau_1}v_1(Z(\tau_1),X_1(\tau_1)) - \mathbb{E}e^{-\alpha\tau_1}S(\tau_1)(1+\delta).
\end{aligned}$$

In the second line we used $v_0(z,x) \geq v_1(z,x) - e^x(1-e^{a-z})(1+\delta)$. Thanks to (36) we have $\mathbb{E}e^{-\alpha\tau_1}v_1(Z(\tau_1),X_1(\tau_1)) \geq \mathbb{E}e^{-\alpha\nu_1}v_1(Z(\nu_1),X_1(\nu_1))$ which leads to

$$\begin{aligned}
v_0(z,x) &\geq \mathbb{E}e^{-\alpha\nu_1}v_1(Z(\nu_1),X_1(\nu_1)) - \mathbb{E}e^{-\alpha\tau_1}S(\tau_1)(1+\delta) \\
&\geq \mathbb{E}e^{-\alpha\nu_1}\left[v_0(Z(\nu_1),X_1(\nu_1)) + S(\nu_1)(1-\delta)\right] - \mathbb{E}e^{-\alpha\tau_1}S(\tau_1)(1+\delta) \\
&= \mathbb{E}e^{-\alpha\nu_1}v_0(Z(\nu_1),X_1(\nu_1)) + \mathbb{E}\left[e^{-\alpha\nu_1}S(\nu_1)(1-\delta) - e^{-\alpha\tau_1}S(\tau_1)(1+\delta)\right].
\end{aligned}$$

Next, note that again because of (36), $\mathbb{E}e^{-\alpha\nu_1}v_0(Z(\nu_1),X_1(\nu_1)) \geq \mathbb{E}e^{-\alpha\tau_2}v_0(Z(\tau_2),X_1(\tau_2))$. Then with $v_0(z,x) \geq v_1(z,x) - e^x(1-e^{a-z})(1+\delta)$ we have

$$\begin{aligned}
v_0(z,x) &\geq \mathbb{E}e^{-\alpha\tau_2}v_0(Z(\tau_2),X_1(\tau_2)) + \mathbb{E}\left[e^{-\alpha\nu_1}S(\nu_1)(1-\delta) - e^{-\alpha\tau_1}S(\tau_1)(1+\delta)\right] \\
&\geq \mathbb{E}e^{-\alpha\tau_2}\left[v_1(Z(\tau_2),X_1(\tau_2)) - S(\tau_2)(1+\delta)\right] \\
&+ \mathbb{E}\left[e^{-\alpha\nu_1}S(\nu_1)(1-\delta) - e^{-\alpha\tau_1}S(\tau_1)(1+\delta)\right] \\
&= \mathbb{E}e^{-\alpha\tau_2}v_1(Z(\tau_2),X_1(\tau_2)) - \mathbb{E}e^{-\alpha\tau_2}S(\tau_2)(1+\delta) \\
&+ \mathbb{E}\left[e^{-\alpha\nu_1}S(\nu_1)(1-\delta) - e^{-\alpha\tau_1}S(\tau_1)(1+\delta)\right].
\end{aligned}$$

Continue this way and recall that $v_0(z,x) \geq 0$ to finally obtain

$$v_0(z,x) \geq \mathbb{E}\sum_{n=1}^{N}\left[e^{-\alpha\nu_n}S(\nu_n)(1-\delta) - e^{-\alpha\tau_n}S(\tau_n)(1+\delta)\right].$$

By sending $N \to \infty$ we get $v_0(z,x) \geq J_0(z,x,\Lambda_0)$ for all possible trading strategies Λ_0. This implies $v_0(z,x) \geq V_0(z,x)$. In a similar way we can show that $v_1(z,x) \geq V_1(z,x)$. In the next part we show the equalities. First recall

$$
\begin{aligned}
&(\alpha - \mathcal{A})v_0(z,x) = 0 \text{ on } (a, z_0^*) \cup (z_0, \infty),\\
&(\alpha - \mathcal{A})v_1(z,x) = 0 \text{ on } (a, z_1),\\
&v_0(z,x) = v_1(z,x) - e^x(1-e^{a-z})(1+\delta) \text{ on } (z_0^*, z_0),\\
&v_1(z,x) = v_0(z,x) + e^x(1-e^{a-z})(1-\delta) \text{ on } (z_1, \infty).
\end{aligned}
$$

Now let τ_1^* and ν_1^* as defined above. Then $\tau_1^* < \infty$ and $\nu_1^* < \infty$ a.s. (see Zhang & Zhang 2008, Lemma 6).Thanks to Dynkin's formula we have

$$
\begin{aligned}
v_0(z,x) &= \mathbb{E}e^{-\alpha\tau_1^*} v_0(Z(\tau_1^*), X_1(\tau_1^*))\\
&= \mathbb{E}e^{-\alpha\tau_1^*}\left[v_1(Z(\tau_1^*), X_1(\tau_1^*)) - S(\tau_1^*)(1+\delta)\right]\\
&= \mathbb{E}e^{-\alpha\tau_1^*} v_1(Z(\tau_1^*), X_1(\tau_1^*)) - \mathbb{E}e^{-\alpha\tau_1^*} S(\tau_1^*)(1+\delta)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}e^{-\alpha\tau_1^*} v_1(Z(\tau_1^*), X_1(\tau_1^*)) &= \mathbb{E}e^{-\alpha\nu_1^*} v_1(Z(\nu_1^*), X_1(\nu_1^*))\\
&= \mathbb{E}e^{-\alpha\nu_1^*}\left[v_0(Z(\nu_1^*), X_1(\nu_1^*)) + S(\nu_1^*)(1-\delta)\right]\\
&= \mathbb{E}e^{-\alpha\nu_1^*} v_0(Z(\nu_1^*), X_1(\nu_1^*)) + \mathbb{E}e^{-\alpha\nu_1^*} S(\nu_1^*)(1-\delta).
\end{aligned}
$$

It follows that

$$
v_0(z,x) = \mathbb{E}e^{-\alpha\nu_1^*} v_0(Z(\nu_1^*), X_1(\nu_1^*)) + \mathbb{E}\left[e^{-\alpha\nu_1^*} S(\nu_1^*)(1-\delta) - e^{-\alpha\tau_1^*} S(\tau_1^*)(1+\delta)\right].
$$

Repeat this process to obtain

$$
v_0(z,x) = \mathbb{E}e^{-\alpha\nu_n^*} v_0(Z(\nu_n^*), X_1(\nu_n^*)) + \mathbb{E}\sum_{k=1}^{n}\left[e^{-\alpha\nu_k^*} S(\nu_k^*)(1-\delta) - e^{-\alpha\tau_k^*} S(\tau_k^*)(1+\delta)\right].
$$

In a similar way we can show

$$
\begin{aligned}
v_1(z,x) &= \mathbb{E}e^{-\alpha\nu_1^*} v_1(Z(\nu_1^*), X_1(\nu_1^*)) \\
&= \mathbb{E}e^{-\alpha\nu_1^*} \left[v_0(Z(\nu_1^*), X_1(\nu_1^*)) + S(\nu_1^*)(1-\delta)\right] \\
&= \mathbb{E}e^{-\alpha\nu_1^*} v_0(Z(\nu_1^*), X_1(\nu_1^*)) + Ee^{-\alpha\nu_1^*} S(\nu_1^*)(1-\delta) \\
&\ \ \vdots \\
&= \mathbb{E}e^{-\alpha\nu_n^*} v_0(Z(\nu_n^*), X_1(\nu_n^*)) + Ee^{-\alpha\nu_n^*} S(\nu_n^*)(1-\delta) \\
&\quad + \mathbb{E}\sum_{k=2}^{n} \left[e^{-\alpha\nu_k^*} S(\nu_k^*)(1-\delta) - e^{-\alpha\tau_k^*} S(\tau_k^*)(1+\delta)\right].
\end{aligned}
$$

In order to complete the proof we have to show that $\mathbb{E}e^{-\alpha\nu_n^*} v_0(Z(\nu_n^*), X_1(\nu_n^*)) \xrightarrow[n\to\infty]{} 0$. Recall that the value function is bounded from above by a constant K_0. Hence

$$
\mathbb{E}e^{-\alpha\nu_n^*} v_0(Z(\nu_n^*), X_1(\nu_n^*)) \leq K_0 \cdot \mathbb{E}e^{-\alpha\nu_n^*}.
$$

Since $\nu_n^* \xrightarrow[n\to\infty]{a.s.} \infty$ and $e^{-\alpha\nu_n^*} \leq 1$ we can apply the dominated convergence theorem and obtain

$$
\lim_{n\to\infty} K_0 \cdot \mathbb{E}e^{-\alpha\nu_n^*} = K_0 \cdot \mathbb{E}\lim_{n\to\infty} e^{-\alpha\nu_n^*} = 0,
$$

and therefore $v_0(z,x) = V_0(z,x)$ and $v_1(z,x) = V_1(z,x)$. $\square$

Although this verification theorem provides sufficient conditions for the triple (z_0^*, z_0, z_1) in order to guarantee the optimality of our results, it is not known if these constants even exist yet if they are unique. However, we will illustrate a solution in a numerical example in the next chapter.

5 A numerical example

We use the following parameters for our model

$$\begin{array}{lll} \alpha = 0.1, & \delta = 0.02, & a = -0.8, \\ \sigma_1 = 0.5, & \sigma_2 = 0.5, & k1 = 2.5, \\ k2 = 12, & \theta_1 = 2, & \theta_2 = 0.7, \\ \theta = 1.3, & \rho = 0.8. & \end{array}$$

The following computations were carried out by using the computer software MATLAB. First we solved equation (25) for z_0^*. Using that, (z_0, z_1) can be obtained by solving equation (29). Finally, C_2 can be obtained from equation (28) and (A_1, A_2), (B_1, B_2) from (21) and (23), respectively. We obtained the following results:

$$\begin{array}{lll} z_0^* = -0.1209, & z_0 = 1.1026, & z_1 = 1.3693 \\ C_2 = -0.2777, & A_1 = 0.0005, & A_2 = 0.8871, \\ B_1 = 0, & B_2 = 0. & \end{array}$$

Furthermore, the obtained parameters satisfy the conditions (35) and (34) given in the verification theorem. Hence one could construct optimal trading strategies in the described way. This illustrates the possible practical use of our computations.

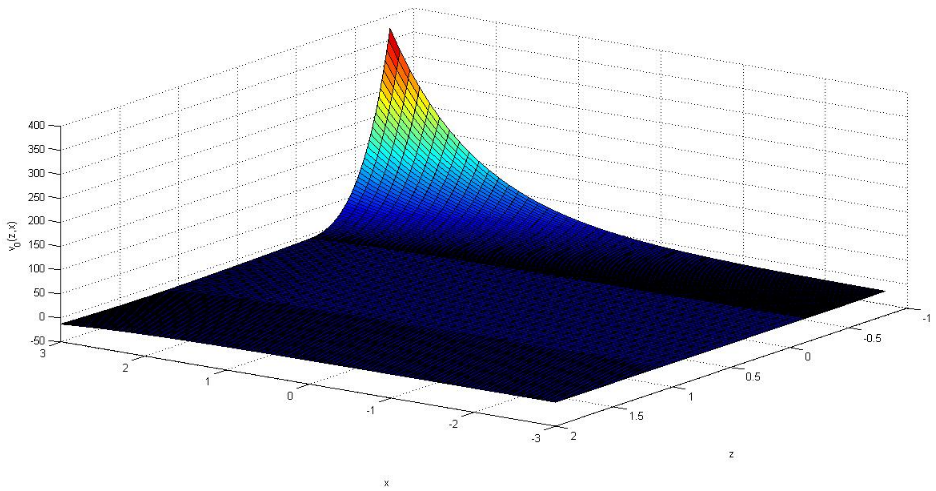

Figure 2: $v_0(z, x)$

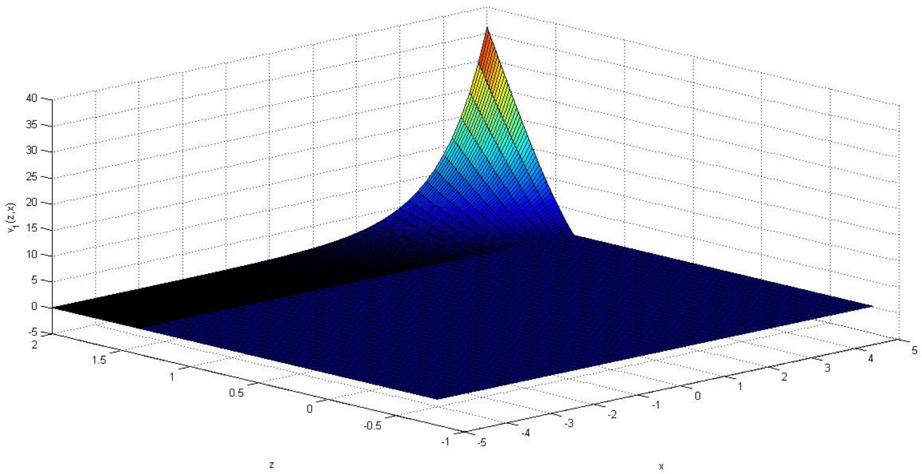

Figure 3: $v_1(z, x)$

6 Conclusion and open problems

The goal of this thesis was to find an optimal trading strategy for a pair of historically correlated stocks. We showed that these optimal trading rules can be constructed by using three threshold levels. If the spread of the asset prices falls below the level z_0 then one should buy the pair. If it reaches the level z_1 after some time, an investor should sell the pair. It was shown that following these rules maximizes the total profit. The three key levels could be attained by solving a set of nonlinear equation. To get these equations we followed the dynamic programming approached and solved the associated Hamilton-Jacobi-Bellman equations by utilizing the smooth-fit method.

Although we were able to prove the optimality of the trading rules constructed with the threshold levels we had to use some assumptions, such as the existence of the inverse of the matrix $\Phi(z)$ for all $z \geq a$ or that v_0 is positive. Nevertheless it was possible to illustrate a numerical example for which a solution (z_0^*, z_0, z_1) existed which satisfied the conditions of the verification theorem. However, it became apparent that with that many parameters for the model it is difficult to monitor the influence of a particular variable on the behavior of the resulting threshold levels.

Therefore it would be interesting to examine these dependencies in more details as well as to be able to prove the existence and uniqueness of the three threshold levels for a set of predetermined model parameters. In this context it is also noteworthy to mention the difficulties arising with the choice of our model settings in contrast to comparable studies such as Song, Q. and Zhang, Q. (2013). The source of most of the difficulties was our choice of the state variable. In taking the differences of the log-prices the price of a position in the pair depended on two variables and hence we had to deal with two-dimensional functions instead of one-dimensional ones as Song and Zhang did by choosing the original difference and therewith the price of the pair as state variable. This fact made not only the computations more complex but also exacerbated the finding of an upper bound for the value functions. With all these difficulties in mind it would be also interesting to find a real life example where one could apply our results.

7 References

[1] Chiu, M. C. and Wong, H. Y. (2011). *Mean-variance portfolio selection of cointegrated assets*. J. Econom. Dynam. Control, 35(8):1369-1385.

[2] Gatev, E., Goetzmann, W. N. and Rouwenhorst, K. G. (2006). *Pairs trading: Performance of a relative-value arbitrage rule*. Review of Finacial Studies, 19(3): 797-827.

[3] Guo, X. and Zhang, Q (2005). *Optimal selling rules in a regime switching model*. IEEE Trans. Automat. Control, 50(9): 1450-1455.

[4] Jurek, J. W. and Yang, H. (2007). *Dynamic portfolio selection in arbitrage*. In EFA 2006 Meetings Paper.

[5] Karatzas, I. and Shreve, S. E. (1991). *Brownian motion and stochastic calculus, volume 113 of Graduate Texts in Mathematics*. Springer-Verlag, New-York, second edition.

[6] Liu, J. and Timmermann, A. (2013). *Optimal convergence trade strategies*. Review of Financial Studies, page hhs130.

[7] Øksendal, B. (2003). *Stochastic differential equations. An introductions with applications*. Universitext. Springer-Verlag, Berlin, sixth edition.

[8] Song, Q. and Zhang, Q. (2013). *An optimal pairs-trading rule*. Automatica J. IFAC, 49(10): 3007-3014.

[9] Tourin, A. and Yan, R. (2013). *Dynamic pairs trading using the stochastic control approach*. J. Econom. Dynam. Control, 37(10):1972-1981.

[10] Touzi, N. (2013). *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE (Fields Institute Monographs)*. Springer-Verlag, New-York.

[11] Zhang, H. and Zhang, Q. (2008). *Trading a mean-reverting asset: buy low and sell high*. Automatica J. IFAC, 44(6): 1511-1518.