Implementation of Pairs Trading Strategies

Øyvind Foshaug

UNIVERSITEIT VAN AMSTERDAM
Faculty of Science
Koortweg- de Vries Institute for Mathematics
Master of Science Thesis

Abstract

In this paper we outline two previously suggested methods for quantitative motivated trading in pairs. We focus on the method of cointegration and a unobserved mean reversion model called the stochastic spread model. The methods are used to implement a search procedure that aims to reveal profitable pairs among all possible pairs available on the German, French and Dutch stock exchanges. The intended user of this application is the trading desk at Amsterdams Effektenkantoor for which this investigation has been done.

Contents

1 Internship assignment

1.1 The company

Amsterdams Effectenkantoor (AEK) operates as an intermediary in transactions in stocks, bonds, real estate shares and derivatives focusing on domestic and foreign institutional investors as well as affluent private clients who appreciate the dedication and service of a personal approach. Within AEK, the research department provides valuable information for clients about movements in financial markets and investigates statistical trading strategies.

1.2 Internship project aim and deliverables

Pairs trading is a trading strategy used to exploit markets that are out of equilibrium assuming that over time they will move to a rational equilibrium. Pairs trading is performed by taking a long position (buying) one security and taking a short position (borrowing) another security. This relative position is called a spread. The profit is made by taking the position in the spread when it is out of equilibrium and unwind the position and make a profit when the spread is at its equilibrium. A spread that has this quality is called mean reverting, with mean indicating equilibrium. AEK is familiar to pairs trading, and pairs trading is carried out on a experimental level based on their industry knowledge. The internship project was initiated to find out if there exist a systematic approach to identify pairs that would make profit for AEK. In order to achieve this it was decided to develop a search procedure for profitable spreads. This search procedure uses available information on securities at stock exchanges in the Netherlands, Germany and France. The search procedure is implemented as as software application with a user interface that give the trader enough flexibility to continuously search for mean reverting spreads and administer the search results. The project was limited to 6 months.

1.3 Developing a search procedure for profitable spreads

Some information on pairs trading is available in the literature and the book by Vidyamurthy (2004) gives the most elaborate discussion on pairs trading. The method of cointegration, Johansen (1991), and the stochastic spread method described in Do, Faff and Hamza (2006) and Elliot, vd Hoek and Malcolm (2004) were chosen as the most promising ones for the search procedure. Johansen (1991) provides an efficient estimation technique of cointegrated vectors given full information of a stochastic processes. The stochastic spread method verifies the existence of a well-known stochastic model for mean reversion called the Ohrnstein-Uhlenbeck process. Statistical tests are used as search criteria for the search procedure to identify a spread as profitable. Furthermore a optimal distance from the equilibrium is calculated in order to give the trader a signal on when to enter the position in the spread. This distance is also used to calculate a historical profit for a given spread.

1.4 Definition of the investment and its return

Pairs trading can be summarized as follows; the trader invests an equal amount in asset A and asset B, $\alpha p_t^A = p_t^B$, provided that it is impossible to buy fraction of assets. This can be done without funding by borrowing a number of shares of assets B, immediately sell these and invest the amount in α shares of asset A. This is called shorting of asset B. We define the investment equation as follows:

$$0 = \log(\alpha) + \log(p_t^A) - \log(p_t^B). \tag{1}$$

The minus sign reflects the fact that asset B is shorted. We express the investment equation in a logarithmic transformation since this will ease the calculation of the returns on the investment. The return on this investment expressed in percentage of the invested amount in asset A over a small horizon $(t-1, t]$ is

$$\log(\frac{\alpha p_t^A}{\alpha p_{t-1}^A}) - \log(\frac{p_t^B}{p_{t-1}^B}) = \log(\frac{p_t^A}{p_t^B}) - \log(\frac{p_{t-1}^A}{p_{t-1}^B}). \tag{2}$$

This is justified by the fact that $\log(s_t/s_{t-1}) \approx (s_t - s_{t-1})/s_{t-1}$. The largest accumulated return is obtained from the largest number of consecutive positive returns. This means that if the behavior of $\log(p_t^A) - \log(p_t^B)$ could be predicted, the investor could enter the trade when this value is low and make profit when then value is high.

1.5 Structure of the thesis

Section 2 describes the method of cointegration and Section 3 the method of stochastic spread. Both methods are described with the aim to develop statistical tests to be used as search criteria, found in section 5. Section 4 describes the optimal trading points and profitability calculations. Section 6 shows the user interface of the software application.

2 Cointegration in time series

This section will focus on the stationary properties of the investment (1) as from Definition 3, any mean reverting process is stationary. Engle and Granger (1987) and Johansen (1988) developed a theoretical framework for a particular class of non stationary $VAR(p)$ processes called cointegrated processes. This class of processes has the property that there may exists stationary linear combinations called cointegrated vectors. If we assume that $\{\log(p_t^A), \log(p_t^B)\}$ in (1) is a non-stationary $VAR(p)$ process, this framework will enable us in some cases to determine γ such that $\log(p_t^A) - \gamma \log(p_t^B)$ is stationary with mean $-\log(\alpha)$. The investment equation will then become

$$0 = \log(\alpha) + \log(p_t^A) - \gamma \log(p_t^B).$$

The value of γ will be determined by the cointegration method, and the long run relationship between the assets determines α. The return on the investment will be

$$\log(\frac{\alpha p_t^A}{\alpha p_{t-1}^A}) - \gamma \log(\frac{p_t^B}{p_{t-1}^B}).$$

The investor is still able to profit from the trade, but the investment may have not have an initial value of 0 and the profit will for this method be dependent on γ. A γ close to zero requires funds to invest in A. A large γ exposes the investor to risk in going short on B. The preferable γ would be around 1, which then again reduces the cointegration method to a test for stationarity in the investment. Section 2.2 characterizes a cointegrated system. Section 2.3 describes maximum likelihood estimation under the restriction of h cointegrated vectors, section 2.4 describes hypothesis testing of h cointegrated vectors using the likelihood ratio test and 2.5 describes the asymptotic distribution of the likelihood ratio test statistics. Section 2.6 describes the estimation of the cointegration vectors.

2.1 Definition and properties of time series

Definition 1. Timeseries. A time series is an infinite sequence

$$y_1, y_2, \dots \tag{3}$$

of random variables or random vectors. These random variables are defined as measurable maps on some underlying probability space. Normally the variables are observed in time and then the index t denotes a real valued date of the observed y.

Definition 2. Covariance stationary time series. The time series y_t is (covariance) stationary if $E(y_t)$ and $E(y_t y_{t+h})$ exists and are finite and do not depend on t, for every $h \in \mathbb{N}$.

Definition 3. Mean reverting time series. A time series is considered mean reverting if

i. The time series is stationary, and

ii. $E(y_t - y_{t-1}|\mathcal{Y}_{t-1}) < 0$ if $y_{t-1} > E(y_t)$
$E(y_t - y_{t-1}|\mathcal{Y}_{t-1}) > 0$ if $y_{t-1} < E(y_t)$,
where $\mathcal{Y}_t$ denotes the history of observation through date t, formally defined as the sigma algebra
$$\mathcal{Y}_t = \sigma\{y_1, y_2, \ldots, y_t\}.$$

2.2 Characterization of cointegration for a VAR system

Consider the $VAR(p)$ equation for $\mathbf{y}_t$ the $(n \times 1)$ collection of random variables

$$\mathbf{y}_t = \alpha + \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2}, \cdots \mathbf{\Phi}_{t-p}\mathbf{y}_{t-p} + \varepsilon_t \tag{4}$$

or

$$\mathbf{\Phi}(L)\mathbf{y}_t = \alpha + \varepsilon_t \tag{5}$$

with

$$\mathbf{\Phi}(L) = [\mathbf{I}_n - \mathbf{\Phi}_1 L - \mathbf{\Phi}_2 L^2 - \cdots - \mathbf{\Phi}_p L^p].$$

We assume

$$E(\varepsilon_t) = 0$$
$$E(\varepsilon_t\varepsilon'_\tau) = \begin{cases} \mathbf{\Omega} & \text{for} \quad t = \tau \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathbf{\Omega}$ is a $(n \times n)$ positive definite matrix. We note that ε_t is a strictly stationary process and it is called a white noise process. We will use the fact that a $VAR(p)$ process is non stationary if the determinantal equation

$$|\mathbf{I}_n - \mathbf{\Phi}_1 z - \mathbf{\Phi}_2 z^2 - \cdots - \mathbf{\Phi}_p z^p| = 0 \tag{6}$$

has at least one solution for z on or inside the unit circle.

Definition 4. A $(n \times 1)$ vector $\mathbf{y}_t$ time series is cointegrated if

i. each of its elements individually are non-stationary and

ii. there exists a nonzero vector $\mathbf{a}'$ such that $\mathbf{a}'\mathbf{y}_t$ is stationary.

The vector $\mathbf{a}$ is not unique. There may be $h < n$ linearly independent vectors $(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_h)$, since if there were n independent vectors, $\mathbf{y}_t$ is itself stationary and the concept of a cointegration will have no meaning. For a cointegrated system, we have that $\mathbf{A}'\mathbf{y}_t$ is stationary where $\mathbf{A}'$ is the following $(h \times n)$ matrix:

$$\mathbf{A}' \equiv \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_h \end{bmatrix}$$

Suppose that $\Delta\mathbf{y}_t = (1 - L)\mathbf{y}_t$ has the Wold representation

$$(1 - L)\mathbf{y}_t = \delta + \mathbf{\Psi}(L)\varepsilon_t, \tag{7}$$

where $\mathbf{\Psi}(L) = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \varepsilon_{t-j}$. The Wold representation is stationary when the determinantal equation in (6) has a single solution for $z = 1$. Necessary conditions for cointegration with cointegrated vectors in $\mathbf{A}$ can be found in Engle and Yoo (1987) or Ogaki and Park (1992). Following these papers, stationarity for $\mathbf{A}'\mathbf{y}_t$ requires that

$$\mathbf{A}'\mathbf{\Psi}(1) = 0 \qquad \text{and} \tag{8}$$

$$\mathbf{A}'\delta = 0. \tag{9}$$

Premultiplying (7) by $\mathbf{\Phi}(L)$ results in :

$$(1 - L)\mathbf{\Phi}(L)\mathbf{y}_t = \mathbf{\Phi}(1)\delta + \mathbf{\Phi}(L)\mathbf{\Psi}(L)\varepsilon_t. \tag{10}$$

Substituting (5) into (10) we have

$$(1 - L)\varepsilon_t = \mathbf{\Phi}(1)\delta + \mathbf{\Phi}(L)\mathbf{\Psi}(L)\varepsilon_t. \tag{11}$$

This relation has to hold for all realizations of ε_t which requires that

$$\mathbf{\Phi}(1)\delta = \mathbf{0} \tag{12}$$

and then

$$(1 - z)\mathbf{I}_n = \mathbf{\Phi}(z)\mathbf{\Psi}(z) \tag{13}$$

for all values of z. For $z = 1$, equation (13) implies that

$$\mathbf{\Phi}(1)\mathbf{\Psi}(1) = \mathbf{0} \tag{14}$$

Let π' denote any row of $\mathbf{\Phi}(1)$. Then (12) and (14) state that $\pi'\mathbf{\Psi}(1) = \mathbf{0}'$ and $\pi'\delta = 0$. By (8) and (9) , π is a cointegrating vector. But by the discussion leading to (7), it must be possible to express π as a linear combination of $\mathbf{a_1}, \ldots, \mathbf{a_h}$, i.e. it exists an $(h \times 1)$ vector $\mathbf{b}$ such that

$$\pi = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_h]\mathbf{b}$$

or

$$\pi' = \mathbf{b}'\mathbf{A}'.$$

Applying this reasoning to each of the rows of $\mathbf{\Phi}(1)$, it follows that there exists an $n \times h, h < n$ matrix $\mathbf{B}$ such that

$$\mathbf{\Phi}(1) = \mathbf{B}\mathbf{A}'. \tag{15}$$

The h linear independent rows of $\mathbf{BA}'$ forms the basis of the space of cointegrating vectors. Note that (14) implies that $\mathbf{\Phi}(1)$ is a singular $(n \times n)$ matrix; linear combinations of the columns of $\mathbf{\Phi}(1)$ of the form $\mathbf{\Phi}(1)\mathbf{x}$ are zero for $\mathbf{x}$ any column of $\mathbf{\Psi}(1)$. Thus the determinant $|\mathbf{\Phi}(z)|$ contains a unit root:

$$|\mathbf{I}_n - \mathbf{\Phi}_1 z - \mathbf{\Phi}_2 z^2 - \cdots - \mathbf{\Phi}_p z^p| = 0 \quad \text{at} \quad z = 1.$$

In the following section we will develop another representation of the matrix $\mathbf{\Phi}(1)$, but with opposite signs. This representation will prove useful in developing the maximum likelihood estimator and the likelihood ratio test.

2.2.1 An alternative representation for the $VAR(p)$ process.

First we consider representation (4):

$$(\mathbf{I}_n - \mathbf{\Phi}_1 L - \mathbf{\Phi}_2 L^2 - \cdots - \mathbf{\Phi}_p L^p)\mathbf{y}_t = \alpha + \varepsilon_t$$

The polynomial on the left side can be rewritten as:

$$\begin{aligned}(\mathbf{I}_n - \mathbf{\Phi}_1 L - &\mathbf{\Phi}_2 L^2 - \cdots - \mathbf{\Phi}_p L^p) \\ &= (\mathbf{I}_n - \rho L) - (\zeta_1 L + \zeta_2 L^2 + \cdots + \zeta_{p-1} L^{p-1})(1 - L),\end{aligned}$$

where

$$\begin{aligned}&\rho \equiv \mathbf{\Phi}_1 + \mathbf{\Phi}_2 + \cdots + \mathbf{\Phi}_p \\ &\zeta_s \equiv -[\mathbf{\Phi}_{s+1} + \mathbf{\Phi}_{s+2} + \cdots + \mathbf{\Phi}_p] \qquad \text{for} \qquad s = 1, 2, \ldots, p-1.\end{aligned}$$

It follows that any $VAR(p)$ process can be written as:

$$(\mathbf{I}_n - \rho L)\mathbf{y}_t - (\zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})(1 - L)\mathbf{y}_t = \alpha + \varepsilon_t$$

or

$$\mathbf{y}_t = \zeta_1 \Delta\mathbf{y}_{t-1} + \zeta_2 \Delta\mathbf{y}_{t-2} + \cdots + \zeta_{p-1} \Delta\mathbf{y}_{t-p+1} + \alpha + \rho \mathbf{y}_{t-1} + \varepsilon_t \tag{16}$$

Subtracting $\mathbf{y}_{t-1}$ from both sides of (16) produces:

$$\Delta\mathbf{y}_t = \zeta_1 \Delta\mathbf{y}_{t-1} + \zeta_2 \Delta\mathbf{y}_{t-2} + \cdots + \zeta_{p-1} \Delta\mathbf{y}_{t-p+1} + \alpha + \zeta_0 \mathbf{y}_{t-1} + \varepsilon_t \tag{17}$$

where

$$\zeta_0 = \rho - \mathbf{I}_n = -(\mathbf{I}_n - \mathbf{\Phi}_1 - \mathbf{\Phi}_2 - \cdots - \mathbf{\Phi}_p) = -\mathbf{\Phi}(1).$$

It follows from (15) that

$$\zeta_0 = -\mathbf{B}\mathbf{A}' \tag{18}$$

also forms the basis of the space of cointegrated vectors.

2.3 Maximum likelihood estimate of ζ_0 under the restriction of h cointegrated vectors.

Consider a sample of $T + p$ observations on $\mathbf{y}_t$, denoted $(\mathbf{y}_{-p+1}, \mathbf{y}_{-p+2}, \cdots, \mathbf{y}_T)$. Assuming Gaussian errors ε_t, the log likelihood of the equation (17) conditional on the p first observations is

$$\begin{aligned}
&\mathcal{L}(\mathbf{\Omega}, \zeta_1, \zeta_2, \cdots, \zeta_{p-1}, \alpha, \zeta_0) \\
&= (-Tn/2)\log(2\pi) - (T/2)\log|\mathbf{\Omega}| \\
&- (1/2)\sum_{t=1}^{T}\Big[(\Delta\mathbf{y}_t - \zeta_1\Delta\mathbf{y}_{t-1} - \zeta_2\Delta\mathbf{y}_{t-2} - \cdots - \zeta_{p-1}\Delta\mathbf{y}_{t-p+1} - \alpha - \zeta_0\mathbf{y}_{t-1})' \\
&\times \mathbf{\Omega}^{-1}(\Delta\mathbf{y}_t - \zeta_1\Delta\mathbf{y}_{t-1} - \zeta_2\Delta\mathbf{y}_{t-2} - \cdots - \zeta_{p-1}\Delta\mathbf{y}_{t-p+1} - \alpha - \zeta_0\mathbf{y}_{t-1})\Big],
\end{aligned} \tag{19}$$

with $|\cdot|$, the determinantal operator. Consider on one hand the OLS regression of the elements of the $(n\times 1)$ vector $\Delta\mathbf{y}_t - \zeta_0\mathbf{y}_{t-1}$ on a constant and $(\Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \cdots, \Delta\mathbf{y}_{t-p+1})$. The noise term vector is

$$[\Delta\mathbf{y}_t - \zeta_0\mathbf{y}_{t-1}] - \Big\{\alpha(\zeta_0) + \zeta_1(\zeta_0)\Delta\mathbf{y}_{t-1} + \zeta_2(\zeta_0)\Delta\mathbf{y}_{t-2} + \cdots + \zeta_{p-1}(\zeta_0)\Delta\mathbf{y}_{t-p+1}\Big\} \tag{20}$$

and has mean zero and is orthogonal to $(\Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \cdots, \Delta\mathbf{y}_{t-p+1})$. Here we treat ζ_0 as given. On the other hand consider the result of $2 \times n$ OLS regressions

$$\Delta\mathbf{y}_t = \pi_0 + \mathbf{\Pi}_1\Delta\mathbf{y}_{t-1} + \mathbf{\Pi}_2\Delta\mathbf{y}_{t-2} + \cdots + \mathbf{\Pi}_{p-1}\Delta\mathbf{y}_{t-p+1} + \mathbf{u}_t \tag{21}$$

from regressing Δy_{it} on $(\Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \cdots, \Delta\mathbf{y}_{t-p+1})$ and

$$\mathbf{y}_{t-1} = \theta + \mathcal{A}_1\Delta\mathbf{y}_{t-1} + \mathcal{A}_2\Delta\mathbf{y}_{t-2} + \cdots + \mathcal{A}_{p-1}\Delta\mathbf{y}_{t-p+1} + \mathbf{v}_t \tag{22}$$

from regressing $y_{i,t-1}$ on $(\mathbf{\Delta y}_{t-1}, \mathbf{\Delta y}_{t-2}, \cdots, \mathbf{\Delta y}_{t-p+1})$. The error terms $\mathbf{u}_t$ and $\mathbf{v}_t$ are orthogonal to $(\mathbf{\Delta y}_{t-1}, \mathbf{\Delta y}_{t-2}, \cdots, \mathbf{\Delta y}_{t-p+1})$ and with mean zero. Therefore

$$\mathbf{u}_t - \zeta_0 \mathbf{v}_t$$

also has these properties. Setting

$$\begin{aligned}
\alpha(\zeta_0) &= \pi_0 - \zeta_0 \theta \\
\zeta_i(\zeta_0) &= \mathbf{\Pi}_i - \zeta_0 \mathcal{A}_i \qquad \text{for } i = 1, 2, \cdots, p-1,
\end{aligned}$$

we can conclude that (20) is equal to:

$$\mathbf{u}_t - \zeta_0 \mathbf{v}_t.$$

Equation (19) becomes for given ζ_0 and $\mathbf{\Omega}$.

$$\begin{aligned}
\mathcal{M}(\mathbf{\Omega}, \zeta_0) &\equiv \mathcal{L}\{\mathbf{\Omega}, \alpha(\zeta_0), \zeta_1(\zeta_0), \cdots, \zeta_{p-1}(\zeta_0), \zeta_0\} \\
&= (-Tn/2)\log(2\pi) - (T/2)\log|\mathbf{\Omega}| - \\
&- (1/2)\sum_{t=1}^{T}[(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)' \mathbf{\Omega}^{-1} (\mathbf{u}_t - \zeta_0 \mathbf{v}_t)].
\end{aligned}$$

If we can find the values ζ_0 and $\mathbf{\Omega}$ for which $\mathcal{M}$ is maximized, then we will find the value that maximizes (19). We can express the estimator of $\mathbf{\Omega}$ in the well known way as

$$\mathbf{\Omega}(\zeta_0) = (1/T)\sum_{t=1}^{T}[(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)'],$$

still treating ζ_0 as given. The likelihood function becomes

$$\begin{aligned}
\mathcal{N}(\zeta_0) &\equiv \mathcal{M}\{\mathbf{\Omega}(\zeta_0), \zeta_0\} \\
&= (-Tn/2)\log(2\pi) - (T/2)\log|\mathbf{\Omega}(\zeta_0)| - (Tn/2) \\
&= (-Tn/2)\log(2\pi) - (Tn/2) \\
&- (T/2)\log\left|(1/T)\sum_{t=1}^{T}[(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)']\right| \qquad (23)
\end{aligned}$$

To maximize $\mathcal{N}$, me must chose ζ_0 that minimizes

$$\left|(1/T)\sum_{t=1}^{T}[(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)(\mathbf{u}_t - \zeta_0 \mathbf{v}_t)']\right| \qquad (24)$$

This will now be our main objective. First we define canonical correlation for random vectors.

Definition 5. We say that the n-dimensional random vectors η_t and ξ_t are in canonical form if

i.

$$E(\eta_t \eta_t') = \mathbf{I}_n \tag{25}$$

$$E(\xi_t \xi_t') = \mathbf{I}_n \tag{26}$$

$$E(\eta_t \xi_t') = \mathbf{R} \tag{27}$$

where

$$\mathbf{R} = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & r_n \end{bmatrix}.$$

ii. The elements of η_t and η_t are ordered in such a way that

$$(1 \geq r_1 \geq r_2 \geq \cdots \geq r_n \geq 0) \tag{28}$$

For now we suppose for simplicity that $\mathbf{u}_t$ and $\mathbf{v}_t$ are in a canonical form as the generalization to (21) and (22) will follow after having obtained some result in this simple case. Let

$$\mathbf{u}_t = \eta_t$$
$$\mathbf{v}_t = \xi_t$$

We would like to choose ζ_0 to minimize the expression of the form (24)

$$\left|(1/T)\sum_{t=1}^{T}[(\eta_t - \zeta_0\xi_t)(\eta_t - \zeta_0\xi_t)']\right| \tag{29}$$

subject to the constraint that $\zeta_0\xi_t$ could make use of only h linear combinations of ξ_t. No restrictions on ζ_0 would mean that (29) would be minimized by OLS regressions of

$$\eta_{it} \quad \text{on} \quad \xi_t$$

for $i = 1, 2, ..., n$. The estimation error $\eta_{it} - \zeta_{0,i}\xi_t$ is uncorrelated with ξ_t :

$$E[(\eta_{it} - \zeta_{0,i}\xi_t)\xi_t'] = 0'. \tag{30}$$

Then

$$\zeta_{0,i} = E(\xi_t \xi_t')^{-1} E(\xi_t \eta_{it}). \tag{31}$$

Conditions (26) and (27) establish that the ith regression will satisfy

$$\left\{E\xi_t\xi_t'\right\}^{-1}\left\{E\xi_t\eta_{it}\right\} = r_i \cdot \mathbf{e}_i,$$

where $\mathbf{e}_i$ denotes the ith column of $\mathbf{I}_n$ Thus, the ith diagonal element in the matrix in (29) becomes:

$$\begin{aligned}
&\left\{E\eta_{it}^2\right\} - \left\{E\eta_{it}\xi_t'\right\}\left\{E\xi_t\xi_t'\right\}^{-1}\left\{E\xi_t\eta_{it}\right\} \\
&= 1 - r_i \cdot \mathbf{e}_i' \cdot \mathbf{I}_n \cdot \mathbf{e}_i \cdot r_i \\
&= 1 - r_i^2
\end{aligned}$$

So for an unrestricted ζ_0 the optimal value for the matrix in (29) would have $1 - r_i^2$ on the diagonal and 0 elsewhere. If we were restricted to only use h linear combinations of ξ_t, the minimal value of (29) will appear by using only the h elements in ξ_t that have the highest correlations. See Johansen (1988) for a reference:

$$\begin{aligned}
&\begin{vmatrix}
1 - r_1^2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 - r_2^2 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 1 - r_h^2 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 1
\end{vmatrix} \\
&= \prod_{i=1}^{h}(1 - r_i^2).
\end{aligned} \tag{32}$$

This method of finding the maximizer for ζ_0 is convenient but assumes canonical error terms. The vectors $\mathbf{u}_t$ and $\mathbf{v}_t$ from (21) and (22) are not canonical. However, we will find matrices $\mathcal{K}$ and $\mathcal{A}$ such that

$$\eta_t = \mathcal{K}'\mathbf{u}_t \tag{33}$$

$$\xi_t = \mathcal{A}'\mathbf{v}_t \tag{34}$$

and η_t and ξ_t are in canonical form. The conditions (25) to (27) becomes:

$$E(\eta_t\eta_t') = \mathcal{K}'\mathbf{\Sigma}_{\mathbf{VV}}\mathcal{K} = \mathbf{I}_n \tag{35}$$

$$E(\xi_t\xi_t') = \mathcal{A}'\mathbf{\Sigma}_{\mathbf{UU}}\mathcal{A} = \mathbf{I}_n \tag{36}$$

$$E(\xi_t\eta_t') = \mathcal{A}'\mathbf{\Sigma}_{\mathbf{UV}}\mathcal{K} = \mathbf{R} \tag{37}$$

where

$$\mathbf{R} = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & r_n \end{bmatrix}.$$

and $\mathbf{\Sigma_{UU}}$ is the covariance matrix of the white noise process $\mathbf{u}_t$ defined in (21), $\mathbf{\Sigma_{UV}}$ is the covariance matrix of $\mathbf{u}_t$ and $\mathbf{v}_t$ and $\mathbf{v}_t$ is the white noise process defined in (22).

We will show that the eigenvalues of two known matrices fully characterizes $\mathcal{K}$ and $\mathcal{A}$. The following proposition must be satisfied.

Proposition 1. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ be the ordered eigenvalues of the matrix

$$\mathbf{\Sigma_{VV}}^{-1}\mathbf{\Sigma_{VU}}\mathbf{\Sigma_{UU}}^{-1}\mathbf{\Sigma_{UV}}.$$

Let $\{\mathbf{k}_1, \cdots, \mathbf{k}_n\}$ be the associated eigenvectors normalized by $\mathbf{k}_i'\mathbf{\Sigma_{VV}}\mathbf{k}_i = 1$. Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq 0$ be the eigenvalues of

$$\mathbf{\Sigma_{UU}}^{-1}\mathbf{\Sigma_{UV}}\mathbf{\Sigma_{VV}}^{-1}\mathbf{\Sigma_{VU}}.$$

with the associated eigenvectors $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ normalized by $\mathbf{a}_i'\mathbf{\Sigma_{UU}}\mathbf{a}_i = 1$. Let

$$\begin{aligned} \mathcal{A} &\equiv [\mathbf{a}_1 \cdots \mathbf{a}_n] \\ \mathcal{K} &\equiv [\mathbf{k}_1 \cdots \mathbf{k}_n]. \end{aligned}$$

Assuming that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct, then

a. $0 \leq \lambda_i < 1$ for $i = 1, 2, \ldots, n$;

b. $\lambda_i = \mu_j$ for $i = 1, 2, \ldots, n$;

c. $\mathcal{K}'\mathbf{\Sigma_{VV}}\mathcal{K} = \mathbf{I}_n$ and $\mathcal{A}'\mathbf{\Sigma_{UU}}\mathcal{A} = \mathbf{I}_n$:

d. $\mathcal{A}'\mathbf{\Sigma_{UV}}\mathcal{K} = \mathbf{R}$, where

$$\mathbf{R^2} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Proof of a. The eigenvalues solves the equation

$$|\lambda \boldsymbol{\Sigma}_{\mathbf{VV}} - \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}}| = 0. \tag{38}$$

If we assume on the contrary that any of these eigenvalues is negative, then since $\boldsymbol{\Sigma}_{\mathbf{VV}}$ is positive definite, $\lambda \boldsymbol{\Sigma}_{\mathbf{VV}}$ is also negative which means that $\boldsymbol{\Sigma}_{\mathbf{VV}} - \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}}$ is negative. Hence the determinant of (38) could not be zero for any value $\lambda < 0$.

□

Proof of c. The eigenvalue-eigenvector pair $(\lambda_i, \mathbf{k}_i)$ satisfies

$$\boldsymbol{\Sigma}_{\mathbf{VV}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i = \lambda_i \mathbf{k}_i. \tag{39}$$

We pre multiply expression (39) by $\mathbf{k}'_j \boldsymbol{\Sigma}_{\mathbf{VV}}$:

$$\mathbf{k}'_j \boldsymbol{\Sigma}_{\mathbf{VV}} \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i = \lambda_i \mathbf{k}'_j \boldsymbol{\Sigma}_{\mathbf{VV}} \mathbf{k}_i \tag{40}$$

Similarly, replace i with j in (39) and pre multiply by $\mathbf{k}'_i \boldsymbol{\Sigma}_{\mathbf{VV}}$:

$$\mathbf{k}'_i \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_j = \lambda_j \mathbf{k}'_i \boldsymbol{\Sigma}_{\mathbf{VV}} \mathbf{k}_j \tag{41}$$

Subtracting (41) from (40) we see that:

$$0 = (\lambda_i - \lambda_j) \mathbf{k}'_j \boldsymbol{\Sigma}_{\mathbf{VV}} \mathbf{k}_i \tag{42}$$

If $i \neq j$ and $\lambda_i \neq \lambda_j$ then (42) establishes that $\mathbf{k}'_j \boldsymbol{\Sigma}_{\mathbf{VV}} \mathbf{k}_i = 0$ for $i \neq j$. Also for $i = j$, the normalization concludes that for distinct eigenvalues, relation (35) holds. Virtually identical calculations show that if we choose the columns of $\mathcal{A}$, the eigenvectors $[\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n]$ such that

$$\boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}} \boldsymbol{\Sigma}_{\mathbf{VV}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{VU}} \mathbf{a}_i = \lambda_i \mathbf{a}_i, \tag{43}$$

condition (36) holds. □

Proof of d. First we transpose (43) and postmultiply by $\boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i$:

$$\mathbf{a}'_i \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}} \boldsymbol{\Sigma}_{\mathbf{VV}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i = \lambda_i \mathbf{a}'_i \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i \tag{44}$$

Similarly, pre multiply (39) by $\mathbf{a}'_i \boldsymbol{\Sigma}_{\mathbf{UV}}$

$$\mathbf{a}'_i \boldsymbol{\Sigma}_{\mathbf{UU}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{UV}} \boldsymbol{\Sigma}_{\mathbf{VV}}{}^{-1} \boldsymbol{\Sigma}_{\mathbf{VU}} \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i = \lambda_j \mathbf{a}'_i \boldsymbol{\Sigma}_{\mathbf{UV}} \mathbf{k}_i \tag{45}$$

Subtracting (45) from (44) results in

$$0 = (\lambda_i - \lambda_j) \mathbf{a}'_i \boldsymbol{\Sigma}_{\mathbf{VV}} \mathbf{k}_j$$

This shows that $\mathbf{a}_i'\mathbf{\Sigma_{VV}}\mathbf{k}_j = 0$ for $\lambda_i \neq \lambda_j$ as required in (37) To find the value of $\mathbf{a}_i'\mathbf{\Sigma_{VV}}\mathbf{k}_j$ for $\lambda_i = \lambda_j$ premultiply (43) by $\mathbf{a}_i'\mathbf{\Sigma_{UV}}$ and using the normalization:

$$\mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathbf{\Sigma_{VV}}^{-1}\mathbf{\Sigma_{VU}}\mathbf{a}_i = \lambda_i \tag{46}$$

For nonsingular $\mathcal{K}$ (35) implies that:

$$\mathbf{\Sigma_{VV}} = [\mathcal{K}]^{-1}\mathcal{K}^{-1} \tag{47}$$

or taking inverses:

$$\mathbf{\Sigma_{VV}}^{-1} = \mathcal{K}\mathcal{K}'$$

Substituting (47) into (46), we find that:

$$\mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathcal{K}\mathcal{K}'\mathbf{\Sigma_{VU}}\mathbf{a}_i = \lambda_i \tag{48}$$

Now

$$\begin{aligned}
\mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathcal{K} &= \mathbf{a}_i'\mathbf{\Sigma_{UV}}[\mathbf{k}_1 \quad \mathbf{k}_2 \quad \cdots \quad \mathbf{k}_n] \\
&= [\mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathbf{k}_1 \quad \cdots \quad \mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathbf{k}_i \quad \cdots \quad \mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathbf{k}_n \quad] \\
&= [0 \quad 0 \quad \cdots \quad \mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathbf{k}_i \quad \cdots \quad 0]
\end{aligned} \tag{49}$$

Substituting (49) into (48), it follows that

$$(\mathbf{a}_i'\mathbf{\Sigma_{UV}}\mathbf{k}_i)^2 = \lambda_i.$$

So r_i from (37) is given by the square root of the eigenvalue λ_i,

$$r_i^2 = \lambda_i. \tag{50}$$

□

Having characterized $\mathcal{K}$ and $\mathcal{A}$, we assume that they are nonsingular. Then equations (33) and (34) allow (24) to be written as

$$\begin{aligned}
&\left|(1/T)\sum_{t=1}^{T}\left[(\mathbf{u}_t - \xi_0\mathbf{v}_t)(\mathbf{u}_t - \xi_0\mathbf{v}_t)'\right]\right| \\
&\quad = \left|(1/T)\sum_{t=1}^{T}\left[[(\mathcal{K}')^{-1}\eta_t - \zeta_0(\mathcal{A}')^{-1}\xi_t][(\mathcal{K}')^{-1}\eta_t - \zeta_0(\mathcal{A}')^{-1}\xi_t]'\right]\right| \\
&\quad = \left|(\mathcal{K}')^{-1}(1/T)\sum_{t=1}^{T}\left[[\eta_t - \mathcal{K}'\zeta_0(\mathcal{A}')^{-1}\xi_t][(\mathcal{K}')^{-1}\eta_t - \mathcal{K}'\zeta_0(\mathcal{A}')^{-1}\xi_t]'\right](\mathcal{K})^{-1}\right| \\
&\quad = |(\mathcal{K}')^{-1}|\left|\sum_{t=1}^{T}\left[[\eta_t - \mathbf{\Pi}\xi_t][\eta_t - \mathbf{\Pi}\xi_t]'\right]\right||(\mathcal{K})^{-1}| \\
&\quad = \left|\sum_{t=1}^{T}\left[[\eta_t - \mathbf{\Pi}\xi_t][\eta_t - \mathbf{\Pi}\xi_t]'\right]\right| \div |\mathcal{K}|^2
\end{aligned}$$

where

$$\mathbf{\Pi} \equiv \mathcal{K}' \zeta_0 (\mathcal{A}')^{-1} \xi_t$$

The method of minimizing (24) is to use the regressors of the first h elements of ξ_t and the value at the optimum is

$$\prod_{i=1}^{h} (1 - r_i^2) \div |\mathcal{K}|^2$$

Moreover, the matrix $\mathcal{K}$ satisfies

$$\mathbf{I}_n = (1/T) \sum_{t=1}^{T} \eta_t \eta_t' = (1/T) \sum_{t=1}^{T} \mathcal{K}' \eta_t \eta_t' \mathcal{K} = \mathcal{K}' \mathbf{\Sigma_{UU}} \mathcal{K}$$

Taking determinants on both sides establish that

$$1/|\mathcal{K}|^2 = |\mathbf{\Sigma_{UU}}|.$$

The optimizing value is then

$$|\mathbf{\Sigma_{UU}}| \times \prod_{i=1}^{h} (1 - r_i^2).$$

The maximum of the log likelihood function of (23) is given by :

$$\mathcal{L} = \mathcal{N}(\zeta_0) = (-Tn/2) \log(2\pi) - (Tn/2) - (T/2) \log |\mathbf{\Sigma_{UU}}| - (T/2) \sum_{i=1}^{h} (1 - \lambda_i), \quad (51)$$

using (50).

2.4 Hypothesis testing

We have seen in that under the restriction of h linear combination of the elements of $\mathbf{y}_t$, the maximum log likelihood is given by (51). This forms the hypothesis H_0: $h < n$ contegrating vectors against the alternative hypothesis H_A: n cointegrating vectors. The maximum log likelihood in the absence of constraints where $h = n$ is:

$$\mathcal{L}_A = (-Tn/2) \log(2\pi) - (Tn/2) - (T/2) \log |\mathbf{\Sigma_{UU}}| - (T/2) \sum_{i=1}^{h} \log(1 - \lambda_i) \quad (52)$$

A likelihood ratio test of H_0 against H_A can be based on

$$\mathcal{L}_A - \mathcal{L}_0 = (-T/2) \sum_{i=h+1}^{n} \log(1 - \lambda_i) \quad (53)$$

To test if the bivariate $VAR(p)$ series described in the introduction to this section contains one cointegration relation means first testing the null hypothesis: $h = 0$ cointegrating relation against the alternative H_A: n cointegrating relations. Then we need to test the null hypothesis of $h = 1$ cointegrating relation against the same alternative as for the first test; if the first test is rejected and the second is accepted, we have statistical evidence of a single cointegrating relation. Johansen (1988) developed the asymptotic properties distribution of the test statistic as we will see in the next section.

2.5 Limiting distribution of the test statistic

This section is only illustrative in the sense that the asymptotic distribution of the test statistic (53) is developed under the restriction that there is no constant term α in (16). In our application this may very well be the case. Without this restriction other critical values will emerge, but the development of the asymptotic distribution is similar. Consider the Wold representation of $\mathbf{\Delta y}_t$ as in (7)

$$\mathbf{\Delta y}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \varepsilon_{t-j}.$$

The null space of $\mathbf{\Psi}' = \sum_{j=0}^{\infty} \mathbf{\Psi'}_j$ given by $\{\xi | \mathbf{\Psi}' \xi = 0\}$ is exactly the range space of ζ_0. We have the two representations

$$\zeta_0 = -\mathbf{BA}' \qquad \text{and} \qquad \mathbf{\Psi}' = \mathbf{EFG}', \tag{54}$$

Where $\mathbf{E}'\mathbf{A} = \mathbf{G}'\mathbf{B} = 0$, and $\mathbf{E}, \mathbf{G}$ are $n \times (n-h)$ and $\mathbf{G}$ is $(n-h) \times (n-h)$, and all three matrices are of full rank. See Granger and Engle (1981) for details on these results.

Definition 6. The following definitions will prove useful:

i. Let $\mathbf{P_B}$ be the projection of $\mathbb{R}^p$ onto the column space spanned by $\mathbf{B}$ with respect to $\mathbf{\Omega}^{-1}$, i.e., $\mathbf{P_B}(\mathbf{\Omega}) = \mathbf{B}(\mathbf{B}'\mathbf{\Omega}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{\Omega}^{-1}$

ii. Let $\mathbf{G}$ be a matrix that fit the description in (54) and such that $\mathbf{GG}' = \mathbf{\Omega}^{-1}(\mathbf{I} - \mathbf{P_B}(\mathbf{\Omega}))$

iii. A result from regression on random coefficients found in Johansen (1984) concludes that $\mathbf{P_B}(\mathbf{\Omega}) = \mathbf{P_B}(\mathbf{\Sigma_{UU}})$

The following proposition derived by Johansen (1988) and Phillips and Durlauf (1986) describing stochastic limit results in terms of stochastic integrals will help us establish the limit distribution.

Proposition 2. Let W be a Brownian motion in n dimensions with covariance matrix $\mathbf{\Omega}$. For $T \to \infty$ it holds that

$$\hat{\mathbf{\Sigma}}_{\mathbf{UU}} \overset{a.s.}{\to} \mathbf{\Sigma}_{\mathbf{UU}}, \tag{55}$$

$$\mathbf{G}'\hat{\mathbf{\Sigma}}_{\mathbf{UV}} \overset{w}{\to} \mathbf{G}' \int_0^1 \mathbf{W}(r) d\mathbf{W}(r)' dr \mathbf{\Psi}' \tag{56}$$

$$\mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{UV}} \overset{a.s.}{\to} \mathbf{A}'\mathbf{\Sigma}_{\mathbf{UV}} \tag{57}$$

$$T^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{VV}} \overset{w}{\to} \mathbf{\Psi}' \int_0^1 \mathbf{W}(r)\mathbf{W}(r)' dr \mathbf{\Psi}' \tag{58}$$

$$\mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{A} \overset{a.s.}{\to} \mathbf{A}'\mathbf{\Sigma}_{\mathbf{VV}}\mathbf{A}'. \tag{59}$$

Consider the eigenvalues that solves the maximum likelihood problem (23)

$$|\hat{\mathbf{\Sigma}}_{\mathbf{VV}}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}} - \lambda \mathbf{I}_n| = 0 \tag{60}$$

Since $\hat{\mathbf{\Sigma}}_{\mathbf{VV}}$ is positive definite, this will be true if and only if

$$|\lambda\hat{\mathbf{\Sigma}}_{\mathbf{VV}} - \hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}| = 0$$

We multiply the matrix inside the determinant in (60) by $(\mathbf{A}, \mathbf{E})$ and its transposed on both sides and obtain:

$$\left|\lambda \begin{bmatrix} \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{A} & \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{E} \\ \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{A} & \mathbf{E}\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{E} \end{bmatrix} - \begin{bmatrix} \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{A} & \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{E} \\ \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{A} & \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{E} \end{bmatrix}\right| = 0 \tag{61}$$

Which is the same as

$$\left|\begin{bmatrix} \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{A}/T & \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{E}/T \\ \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{A}/T & \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VV}}\mathbf{E}/T \end{bmatrix} - \mu \begin{bmatrix} \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{A} & \mathbf{A}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{E} \\ \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{A} & \mathbf{E}'\hat{\mathbf{\Sigma}}_{\mathbf{VU}}\hat{\mathbf{\Sigma}}_{\mathbf{UU}}{}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{E} \end{bmatrix}\right| = 0 \tag{62}$$

where

$$\hat{\mu}_1 = (T\hat{\lambda}_n)^{-1}, \ldots, \hat{\mu}_n = (T\hat{\lambda}_1)^{-1}$$

In this way the distributions of largest μ's are given as continuous functions of the ordered eigenvalues of (61). From (57), $\hat{\mathbf{\Sigma}}_{\mathbf{UV}}\mathbf{E} \overset{a.s}{\to} \mathbf{\Sigma}_{\mathbf{UV}}\mathbf{E} \equiv \mathbf{Q}$ and (62) converges in distribution to the ordered eigenvalues $1/\mu_1, \ldots, 1/\mu_n$ of the equation

$$\left|\begin{bmatrix} 0 & 0 \\ 0 & \mathbf{E}'\mathbf{\Psi}\int_0^1 \mathbf{W}(r)\mathbf{W}'(r)dr\mathbf{\Psi}'\mathbf{E} \end{bmatrix} - \mu \begin{bmatrix} \mathbf{A}'\mathbf{\Sigma}_{\mathbf{UV}}\mathbf{\Sigma}_{\mathbf{UU}}{}^{-1}\mathbf{\Sigma}_{\mathbf{UV}}\mathbf{A} & \mathbf{A}'\mathbf{\Sigma}_{\mathbf{UV}}\mathbf{\Sigma}_{\mathbf{UU}}{}^{-1}\mathbf{Q} \\ \mathbf{Q}'\mathbf{\Sigma}_{\mathbf{UU}}{}^{-1}\mathbf{\Sigma}_{\mathbf{UV}}\mathbf{A} & \mathbf{Q}'\mathbf{\Sigma}_{\mathbf{UU}}{}^{-1}\mathbf{Q} \end{bmatrix}\right| = 0$$

for which the left side is identical to

$$
|\mu \mathbf{A}'\mathbf{\Sigma_{UV}}\mathbf{\Sigma_{UU}}^{-1}\mathbf{\Sigma_{UV}}\mathbf{A}|\Big|\mathbf{E}'\mathbf{\Psi}\int_0^1 \mathbf{W}(r)\mathbf{W}'(r)dr\mathbf{\Psi}'\mathbf{E} - \mu\mathbf{Q}'\Big[\mathbf{\Sigma_{UU}}^{-1} \\
- \mathbf{\Sigma_{UU}}^{-1}\mathbf{\Sigma_{UV}}\mathbf{A}(\mathbf{A}'\mathbf{\Sigma_{UV}}\mathbf{\Sigma_{UU}}^{-1}\mathbf{\Sigma_{UV}}\mathbf{A})^{-1}\mathbf{A}'\mathbf{\Sigma_{UV}}\mathbf{\Sigma_{UU}}^{-1}\Big]\mathbf{Q}\Big|. \tag{63}
$$

Now we use definition (6). The term in the square brackets is then equal to

$$
\mathbf{\Sigma_{UU}}^{-1}(\mathbf{I} - \mathbf{P_B}(\mathbf{\Sigma_{UU}})) = \mathbf{\Omega}^{-1}(\mathbf{I} - \mathbf{P_B}(\mathbf{\Omega})) = \mathbf{GG}'. \tag{64}
$$

Also using (56) on the second term, equation (63) becomes, when eliminating the first factor,

$$
\left|\mathbf{E}'\mathbf{\Psi}\int_0^1 \mathbf{W}(r)\mathbf{W}(r)'dr\mathbf{\Psi}'\mathbf{E} - \mu\mathbf{E}'\mathbf{\Psi}\int_0^1 \mathbf{W}(r)d\mathbf{W}(r)'\mathbf{GG}'\int_0^1 \mathbf{W}(r)d\mathbf{W}(r)'\mathbf{\Psi}'\mathbf{G}\right| = 0. \tag{65}
$$

We rewrite this equation using the second relation in (54), $|\mathbf{FF}'| \neq 0$ and $|\mathbf{D}| \neq 0$. We then get

$$
\left|\mathbf{G}'\int_0^1 \mathbf{W}(r)\mathbf{W}'(r)dr\mathbf{G} - \mu\mathbf{G}'\int_0^1 \mathbf{W}(r)d\mathbf{W}(r)'\mathbf{GG}'\int_0^1 \mathbf{W}(r)d\mathbf{W}(r)'\mathbf{G}\right| = 0.
$$

But $\mathbf{B} = \mathbf{G}'\mathbf{W}$ is also a Brownian motion and has variance $\mathbf{G}'\mathbf{\Omega}\mathbf{G} = \mathbf{I}$. So the we can conclude that the eigenvalues $T\hat{\lambda}_{h+1}, \dots, T\hat{\lambda}_n$ converge in distribution to the eigenvalues that solves the equation

$$
\left|\lambda\int_0^1 \mathbf{W}(r)\mathbf{W}'(r)dr - \int_0^1 \mathbf{W}(r)d\mathbf{W}(r)'\int_0^1 d\mathbf{W}(r)\mathbf{W}(r)'\right| = 0.
$$

as $T \to \infty$. Then the sum of the eigenvalues $\hat{\lambda}_{h+1}, \dots, \hat{\lambda}_n$ is given by:

$$
tr\left\{\int_0^1 d\mathbf{W}(r)\mathbf{W}(r)'\left(\int_0^1 \mathbf{W}(r)\mathbf{W}'(r)dr\right)^{-1}\int_0^1 \mathbf{W}(r)d\mathbf{W}(r)'\right\}. \tag{66}
$$

Now consider (53) :

$$
\mathcal{L}_A - \mathcal{L}_0 = (-T/2)\sum_{i=h+1}^{n}\log(1 - \lambda_i).
$$

We rewrite and obtain

$$
2(\mathcal{L}_A - \mathcal{L}_0) = (-T)\sum_{i=h+1}^{n}\log(1 - \lambda_i) = \sum_{i=h+1}^{n} T\lambda_i + o_p(1).
$$

We conclude that this expression converge in distribution to (66) and the critical values for this statistic can be found in Johansen (1988).

2.6 Estimating the cointegrating vectors

The log likelihood function is maximized by selecting as regressors the first h elements of ξ_t. Since from (34) $\xi_t = \mathcal{A}'\mathbf{v}_t$, this means using $\mathbf{A}'\mathbf{v}_t$ as regressors, where the $(n \times h)$ matrix $\mathbf{A}$ denotes the first h columns of the $(n \times n)$ matrix $\mathcal{A}$. Thus

$$\zeta_0 \mathbf{v}_t = -\mathbf{B}\mathbf{A}'\mathbf{v}_t \tag{67}$$

The matrix $\mathbf{A}$ is found using the normalized eigenvectors from (43). Let $\mathbf{w} \equiv \mathbf{A}'\mathbf{v}_t$ we have that the value of $\mathbf{B}$ that maximizes the likelihood is given by regressing $\mathbf{u}_t$ on $\mathbf{w}_t$, obtaining the OLS estimate of $\mathbf{B}$ as

$$\mathbf{B} = -\Big[(1/T)\sum_{t=1}^{T} \mathbf{u}_t\mathbf{w}_t'\Big]\Big[(1/T)\sum_{t=1}^{T} \mathbf{w}_t\mathbf{w}_t'\Big] \tag{68}$$

But $\mathbf{w}_t$ is composed of h canonical covariates from (35) so

$$\Big[(1/T)\sum_{t=1}^{T} \mathbf{w}_t\mathbf{w}_t'\Big] = \mathbf{I}_h \tag{69}$$

Also

$$\begin{aligned}\Big[(1/T)\sum_{t=1}^{T} \mathbf{u}_t\mathbf{w}_t'\Big] &= \Big[(1/T)\sum_{t=1}^{T} \mathbf{u}_t\mathbf{v}_t'\mathbf{A}\Big] \\ &= \boldsymbol{\Sigma}_{\mathbf{UV}}\mathbf{A}\end{aligned} \tag{70}$$

We get by substituting (69) and (70) into (68)

$$\mathbf{B} = -\boldsymbol{\Sigma}_{\mathbf{UV}}\mathbf{A},$$

and so, from (18), ζ_0 is given by

$$\zeta_0 = \boldsymbol{\Sigma}_{\mathbf{UV}}\mathbf{A}\mathbf{A}'.$$

3 The hidden Ohrnstein-Uhlenbeck model

In this section we will formulate a model for the investment in equation (1). The model assumes that the true mean reverting process can be described as a discrete version of the Ohrstein-Uhlenbeck continuous stochastic process. Elliot, Hoek and Malcolm (2004) suggest to use this process to model mean reverting pairs and Do, Faff and Hamza (2006) carry this idea further by combining it with the

Arbitrage Pricing Theory, Ross (1976). They claim that if the two stocks that form a pair belong to the same industry sector and are influenced by the same set of risk factors, the APT-theory will explain the difference between the prices returns, leaving the 'true' mean reversion to be modeled by the OU-process. The APT operates on return of assets instead of the prices themselves as in (1), and the investment expressed in returns becomes

$$\log(\alpha r_t^A) - \log(r_t^B) - \omega \log(r_t^X) = 0$$

The first two terms are explained by the OU-process and the third by APT. Let r_t^A be the logarithmic return of asset A and likewise for asset B. Let r_t^X be the excess return of a single risk factor described by the APT. Then we can derive $\log(p_t^A) = \log(p_0^A) + \sum_{i=0}^{i=t} \log(r_i^A)$ and likewise for X and B. The investment (1) becomes

$$\log(\alpha p_t^A) - \log(p_t^B) - \omega \log(p_t^X) = 0.$$

The value of ω is determined by the stochastic spread method. Depending on the sign of the estimated ω, the investor must invest or borrow an asset reflecting the index with this portion of its value. Both papers mentioned above assume that the OU-process is observed only with model noise and uses a state-space model to represent this as a dynamical system, described in section 3.3. The state-space representation is a form that can be analyzed with the Kalman filter introduced by R.E. Kalman (1960). Section 3.3.2 describes the maximum likelihood estimates of the state-space representation using the Kalman filter. The unobserved OU-process, i.e. the process path with highest probability given the observed data, is reconstructed using the Kalman smoothing procedure described in 3.3.1.
Section 3.1 describes the Ohrnstein-Uhlenbeck process, the following section 3.2 describes the APT theory and how it is applied to mean reverting pairs.

3.1 Ornstein-Uhlenbeck process

In this section we consider the one dimensional stochastic process given by the stochastic differential equation (SDE)

$$dr_t = -\theta(r_t - \mu)dt + \sigma dWt, \qquad \theta > 0 \tag{71}$$

known as the Ornstein-Uhlenbeck continuous time process. We will show that there exists a stationary solution to this SDE that could be used to model our spread. For the purpose of intuition about mean reversion, we compare (71) to the geometric Brownian motion SDE

$$dS_t = \alpha S_t dt + \sigma S_t dWt \tag{72}$$

with drift α and volatility σ. For the Ornstein-Uhlenbeck process, the drift is negative for $r_t > \mu$ and positive for $r_t > \mu$, causing the process to revert to the mean μ with speed θ. Applying Ito's lemma to the function $f(r_t, t) = r_t e^{\theta t}$, we get

$$\begin{aligned} df(r_t, t) &= \theta r_t e^{\theta t} dt + e^{\theta t} dr_t \\ &= e^{\theta t} \theta \mu dt + \sigma e^{\theta t} dW_t. \end{aligned}$$

Integrating from 0 to t, we get

$$r_t e^{\theta t} = r_0 + \int_0^t e^{\theta s} \theta \mu ds + \int_0^t \sigma e^{\theta s} dW_s \tag{73}$$

And so

$$r_t = r_0 e^{-\theta t} + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{\theta(s-t)} dW_s \tag{74}$$

Firstly we see

$$E(r_t) = r_0 e^{-\theta t} + \mu(1 - e^{-\theta t}).$$

Consequently,

$$\lim_{t \to \infty} E(r_t) = \mu$$

We let $s \wedge t = min(s, t)$ and we get

$$\begin{aligned} cov(r_t, r_s) &= E[(r_s - E[r_s])(r_t - E[r_t])] \\ &= E[\int_0^t \sigma e^{\theta(u-s)} dW_u \int_0^s \sigma e^{\theta(v-t)} dW_v] \\ &= \sigma^2 e^{-\theta(t+s)} E[\int_0^s e^{\theta u} dW_u \int_0^t e^{\theta v} dW_v] \\ &= \frac{\sigma^2}{2\theta} e^{-\theta(t+s)} (e^{2\theta(s \wedge t)} - 1), \end{aligned} \tag{75}$$

using the Ito isometry in the second last line of the calculation. Setting $s = t$, we get:

$$VAR(r_t) = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}).$$

Consequently,

$$\lim_{t \to \infty} VAR(r_t) = \frac{\sigma^2}{2\theta}.$$

If we set $r_0 \sim N(\mu, \frac{\sigma^2}{2\theta})$, the solution in (74) is stationary according to Definition 2. We rewrite equation (74) with $\Delta t = t_i - t_{i-1}$, keeping Δt small to obtain the discrete version of this process:

$$r_t = \mu(1 - e^{-\theta\Delta t}) + e^{-\theta\Delta t} r_{t-1} + \varepsilon_t, \tag{76}$$

with ε_t a white noise process. This version is necessary for applying the Kalman filter as we will see in the following sections.

3.2 Asset pricing theory

Ross (1976) assumed that the return of a single asset can be formulated as

$$a_t = r_t + \beta \mathbf{b}_t + \varepsilon_t, \tag{77}$$

where r_t is the deterministic risk free return, $\mathbf{b}_t = [(f_t^1 - r_t), (f_t^2 - r_t), \ldots, (f_t^n - r_t)]$ is the return above the risk free return from some factors $f^1, \ldots, f^n$ and the prediction error ε_t that is uncorrelated with a_t . We assume that the these factors have mean zero and that their variance exists. The APT says that the expected return is linear in the factor weights:

$$E(a_t) = r_t + \beta E(\mathbf{b}_t). \tag{78}$$

with $\beta = [\beta_1, \beta_1, \ldots \beta_n]$. We apply this theory to the difference between two assets returns a_t, b_t which then can be formulated as

$$a_t - b_t = \alpha \mathbf{b}_t + \tilde{\varepsilon}_t,$$

with $\alpha = [(\beta_1^a - \beta_1^b), \ldots, (\beta_n^a - \beta_n^b)]'$. We assume that $\tilde{\varepsilon}_t$ is independent of ε_t and that two assets can be explained by the same factors $f^1, \ldots, f^n$. In our numerical experiments we will use an industry index as a single factor to this purpose, explaining two assets in the same industry.

3.3 The state-space representation of a dynamic system

A model of the observed variable y can be described in terms of the possibly unobserved ξ_t known as the state variable. We consider the following definition:

Definition 7. The state-space representation of the dynamics of y is given by the following system of equations.

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1} \tag{79}$$

$$y_t = \mathrm{A}x_t + \mathbf{H}'\xi_t + w_t, \tag{80}$$

where the matrices **F** and **H** are assumed to be known of dimensions (2×2) and (2×1), and A is a scalar. The errors $\mathbf{v}_t$ and w_t are white noise processes with assumed known parameters **Q** and R :

$$E(\mathbf{v}_t \mathbf{v}'_\tau) = \begin{cases} \mathbf{Q} & \text{for} \quad t = \tau \\ 0 & \text{otherwise} \end{cases} \tag{81}$$

$$E(w_t w_\tau) = \begin{cases} \mathrm{R} & \text{for} \quad t = \tau \\ 0 & \text{otherwise} \end{cases}, \tag{82}$$

and are assumed to be uncorrelated at all lags:

$$E(\mathbf{v}_t w_\tau) = 0 \qquad \text{for all } t \text{ and } \tau. \tag{83}$$

We also assume that x_t is predetermined which implies that x_t is independent of ξ_{t+1} and w_{t+s} for $s = 0, 1, 2, \ldots$.

We will keep in mind that the discretized OU-process of equation (76) is here represented as the state variable with the constant term set to $\xi_{t,1}$, and the predetermined variable x_t represent the risk factor excess return described in section 3.2. The system of equations (79) through (83) is typically used to describe a finite series of observations $\{y_1, y_2, \ldots, y_T\}$. Initial conditions are needed for ξ_1:

$$E(\mathbf{v}_t \xi'_1) = 0 \qquad \text{for all } t \text{ and } \tau. \tag{84}$$

$$E(w_t \xi'_1) = 0 \qquad \text{for all } t \text{ and } \tau. \tag{85}$$

The state equation (79) implies that ξ_t can be written as a linear function of $(\xi_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_t)$:

$$\begin{aligned} \xi_t = \mathbf{v}_t + \mathbf{F}\mathbf{v}_{t-1} + \mathbf{F}^2\mathbf{v}_{t-2} + \cdots + \mathbf{F}^{t-2}\mathbf{v}_2 + \mathbf{F}^{t-1}\xi_1 \\ \text{for } t = 2, 3, \ldots, T. \end{aligned} \tag{86}$$

Thus (84) and (81) imply that $\mathbf{v}_t$ is uncorrelated with lagged values of ξ:

$$E(\mathbf{v}_t \xi'_\tau) = 0 \qquad \text{for } t = t-1, t-2, \ldots, 1. \tag{87}$$

Similarily

$$E(w_t \xi'_\tau) = 0 \qquad \text{for } \tau = t-1, t-2, \ldots, 1.$$

3.3.1 Derivation of the Kalman Filter

This section aims at deriving the Kalman filter or the linear least square forecast of the state vector on the basis of data observed through date t:

$$\hat{\xi}_{t+1|t} \equiv \hat{E}(\xi_{t+1} | \mathcal{Y}_t),$$

where

$$\mathcal{Y}_t \equiv (y_t, y_{t-1}, \cdots, y_1, x_t, x_{t-1}, \ldots, x_1)$$

and $\hat{E}(\xi_{t+1}|\mathcal{Y}_t)$ is the linear projection on $\mathcal{Y}_t$ and a constant. The Kalman procedure calculates these forecasts recursively, generating $\hat{\xi}_{1|0}, \hat{\xi}_{2|1}, \ldots, \hat{\xi}_{T|T-1}$ in succession, thus improving the estimate with the added information. Associated with each of these forecasts is the mean squared error (MSE)

$$\mathbf{P}_{t+1|t} \equiv E[(\xi_{t+1} - \hat{\xi}_{t+1|t})(\xi_{t+1} - \hat{\xi}_{t+1|t})']. \tag{88}$$

The recursion begins with $\hat{\xi}_{1|0}$ which is the unconditional mean of ξ_1,

$$\hat{\xi}_{1|0} = E(\xi_1)$$

Since by (86), ξ_t is stationary and the unconditional mean of ξ_t can be found by taking expectations of (79) producing

$$E(\xi_{t+1}) = \mathbf{F}E(\xi_t).$$

The matrix $\mathbf{F}$ has no singular eigenvalues if we assume stationarity in equation (79). This equation therefore has one unique solution in $E(\xi_t) = 0$. Similarly the unconditional variance $\mathbf{P}$ can be found by:

$$E[(\xi_{t+1}\xi'_{t+1})] = E[(\mathbf{F}\xi_t + \mathbf{v}_{t+1})(\xi'_t\mathbf{F}' + \mathbf{v}'_{t+1})] = \mathbf{F}E[\xi_t\xi'_t]\mathbf{F}' + E[\mathbf{v}_{t+1}\mathbf{v}'_{t+1}].$$

The cross product has disappeared in the light of (87). Then

$$\mathbf{P}_{1|0} = \mathbf{F}\mathbf{P}_{1|0}\mathbf{F}' + \mathbf{Q}.$$

Given the values of $\hat{\xi}_{1|0}$ and $\mathbf{P}_{1|0}$ the next step is to calculate the magnitudes for the following date, $\hat{\xi}_{2|1}$ and $\mathbf{P}_{2|1}$. Since we have assumed that x_t contains no information on ξ_t beyond that contained in $\mathcal{Y}_{t-1}$,

$$\hat{E}(\xi_t|x_t, \mathcal{Y}_{t-1}) = \hat{E}(\xi_t|\mathcal{Y}_{t-1}) = \hat{\xi}_{t|t-1}.$$

Next consider forecasting the value of y_t:

$$\hat{y}_{t|t-1} \equiv \hat{E}(\xi_t|x_t, \mathcal{Y}_{t-1}).$$

From (80) we have that

$$\hat{E}(\xi_t|x_t, \mathcal{Y}_{t-1}) = \mathrm{A}x_t + \mathbf{H}'\xi_t,$$

and so, from the law of iterated projections,

$$\hat{y}_{t|t-1} = \mathrm{A}x_t + \mathbf{H}'\hat{E}(\xi_t|x_t, \mathcal{Y}_t) = \mathrm{A}x_t + \mathbf{H}'\hat{\xi}_{t|t-1}. \tag{89}$$

Its forecast error is

$$y_t - \hat{y}_{t|t-1} = \mathrm{A}x_t + \mathbf{H}'\xi_t + w_t - \mathrm{A}x_t - \mathbf{H}'\hat{\xi}_{t|t-1} = \mathbf{H}(\xi_t - \hat{\xi}_{t|t-1}) + w_t.$$

with MSE

$$E[(y_t - \hat{y}_{t|t-1})^2] = E[\mathbf{H}'(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})\mathbf{H}] + E(w_t^2). \tag{90}$$

Cross-product terms have disappeared, since

$$E[w_t(\xi_t - \hat{\xi}_{t|t-1})'] = \mathbf{0} \tag{91}$$

To justify this, recall from by (85) that w_t and ξ_t are uncorrelated and $\hat{\xi}_{t|t-1}$ is a linear function of $\mathcal{Y}_{t-1}$ so it too must be uncorrelated with w_t. Using (82) and (88), (90) can be written

$$E[(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})'] = \mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}' + \mathrm{R}. \tag{92}$$

The inference about the current value of ξ_t based on the observation of y_t is given by:

$$\hat{\xi}_{t|t} = \hat{E}(\xi_t | y_t, x_t, \mathcal{Y}_{t-1}) = \hat{E}(\xi_t | y_t)$$

This is evaluated using the formula for updating a linear projection as described in Hamilton (1994):

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + E[(\xi_t - \hat{\xi}_{t|t-1})(y_t - \hat{y}_{t|t-1})] \times E[(y_t - \hat{y}_{t|t-1})^2]^{-1} \times (y_t - \hat{y}_{t|t-1}). \tag{93}$$

But

$$\begin{aligned} &E\{(\xi_t - \hat{\xi}_{t|t-1})(y_t - \hat{y}_{t|t-1})\} \\ &\quad = E\{[\xi_t - \hat{\xi}_{t|t-1}][\mathbf{H}'(\xi_t - \hat{\xi}_{t|t-1}) + w_t]\} \\ &\quad = E[(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})'\mathbf{H}] \\ &\quad = \mathbf{P}_{t|t-1}\mathbf{H} \end{aligned} \tag{94}$$

using (91) and (88). Substituting (94), (92) and (89) into (93) gives

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{H}'(\mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H} + \mathrm{R})^{-1}(y_t - \mathrm{A}x_t - \mathbf{H}'\hat{\xi}_{t|t-1}). \tag{95}$$

Its MSE is:

$$\begin{aligned} \mathbf{P}_{t|t} \equiv\ & E[(\xi_t - \hat{\xi}_{t|t})(\xi_t - \hat{\xi}_{t|t})'] = E[(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})'] \\ & - \{E[(\xi_t - \hat{\xi}_{t|t-1})(y_t - \hat{y}_{t|t-1})]\} \\ & \times E[(y_t - \hat{y}_{t|t-1})^2]^{-1} \\ & \times \{E[(y_t - \hat{y}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})']\} \\ & = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{H}'(\mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H} + \mathrm{R})^{-1}\mathbf{H}'\mathbf{P}_{t|t-1} \end{aligned} \tag{96}$$

The state equation (79) is used to forecast ξ_{t+1}:

$$\begin{aligned}\hat{\xi}_{t+1|t} &= \hat{E}(\xi_{t+1}|\mathcal{Y}_t)\\ &= \mathbf{F}\hat{E}(\xi_t|\mathcal{Y}_t) + \hat{E}(\mathbf{v}_t|\mathcal{Y}_t)\\ &= \mathbf{F}\hat{\xi}_{t|t} + \mathbf{0}.\end{aligned} \tag{97}$$

Substituting (95) into (97),

$$\begin{aligned}\hat{\xi}_{t+1|t} &= \mathbf{F}\hat{\xi}_{t|t-1}\\ &\quad + \mathbf{F}\mathbf{P}_{t|t-1}\mathbf{H}(\mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H} + \mathrm{R})^{-1}(y_t - \mathrm{A}x_t - \mathbf{H}'\hat{\xi}_{t|t-1}).\end{aligned}$$

Its MSE is found from (97) and (79) :

$$\begin{aligned}\mathbf{P}_{t+1|t} &= E[(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})']\\ &= E[(\mathbf{F}\xi_t + \mathbf{v}_t - \mathbf{F}\hat{\xi}_{t|t})(\mathbf{F}\xi_t + \mathbf{v}_t - \mathbf{F}\hat{\xi}_{t|t})']\\ &= \mathbf{F}E[(\xi_t - \hat{\xi}_{t|t})(\xi_t - \hat{\xi}_{t|t})']\mathbf{F}' + E[\mathbf{v}_{t+1}\mathbf{v}'_{t+1}]\\ &= \mathbf{F}\mathbf{P}_{t|t}\mathbf{F}' + \mathbf{Q},\end{aligned} \tag{98}$$

with cross-products again clearly zero. Substituting (96) into (98) produces:

$$\mathbf{P}_{t+1|t} = \mathbf{F}[\mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{H}(\mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}' + \mathrm{R})^{-1}\mathbf{H}'\mathbf{P}_{t|t-1}]\mathbf{F} + \mathbf{Q}.$$

This completes one iteration step.

3.3.2 Kalman smoothing estimates

Suppose we were told the true value of ξ_{t+1} at t. Then the estimate of ξ_t can be expressed as follows:

$$\begin{aligned}\hat{E}(\xi_t|\xi_{t+1}, \mathcal{Y}_t) = \quad & \hat{\xi}_{t|t} + \{E[(\xi_t - \hat{\xi}_{t|t})(\xi_{t+1} - \hat{\xi}_{t|t})']\}\\ & \times\{E[(\xi_{t+1} - \hat{\xi}_{t+1|t})(\xi_{t+1} - \hat{\xi}_{t|t})']\}^{-1}\\ & \times(\xi_{t+1} - \hat{\xi}_{t+1|t}),\end{aligned} \tag{99}$$

by the law for updating a linear projection. Using (79) and (97), the first term in the product on the right side of (99) can be written

$$E[(\xi_t - \hat{\xi}_{t|t})(\xi_{t+1} - \hat{\xi}_{t|t})'] = E[(\xi_t - \hat{\xi}_{t|t})(\mathbf{F}\xi_t + v_{t+1} - \mathbf{F}\hat{\xi}_{t|t})'].$$

Since v_{t+1} is uncorrelated with ξ_t and $\hat{\xi}_{t|t}$, we have:

$$E[(\xi_t - \hat{\xi}_{t|t})(\xi_{t+1} - \hat{\xi}_{t|t})'] = E[(\xi_t - \hat{\xi}_{t|t})(\xi_t - \hat{\xi}_{t|t})'\mathbf{F}'] = \mathbf{P}_{t|t}\mathbf{F}' \tag{100}$$

Substituting (100) and (88) into (99) produces:

$$\hat{E}(\xi_t|\xi_{t+1}, \mathcal{Y}_t) = \hat{\xi}_{t|t} + \mathbf{P}_{t|t}\mathbf{F}'\mathbf{P}_{t+1|t}^{-1}(\xi_{t+1} - \hat{\xi}_{t+1|t})$$

Defining

$$\mathbf{J}_t \equiv \mathbf{P}_{t|t}\mathbf{F}'\mathbf{P}_{t+1|t}^{-1}, \tag{101}$$

we have

$$\hat{E}(\xi_t|\xi_{t+1}, \mathcal{Y}_t) = \hat{\xi}_{t|t} + \mathbf{J}_t(\xi_{t+1} - \hat{\xi}_{t+1|t})$$

We claim that

$$\hat{E}(\xi_t|\xi_{t+1}, \mathcal{Y}_t) = \hat{E}(\xi_t|\xi_{t+1}, \mathcal{Y}_T)$$

To verify this it is sufficient to use the fact that the error $\xi_{t+1} - \hat{E}(\xi_t|\xi_{t+1}, \mathcal{Y}_t)$ is uncorrelated with $\mathbf{y}_{t+j}$ and $\mathbf{x}_{t+j}$ for $j > 0$. Taking conditional expectation w.r.t $\mathcal{Y}_T$, we get

$$\hat{E}(\xi_t|\mathcal{Y}_T) = \hat{\xi}_{t|t} + \mathbf{J}_t(\hat{E}(\xi_{t+1}|\mathcal{Y}_T) - \hat{\xi}_{t+1|t}),$$

or

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} + \mathbf{J}_t(\hat{\xi}_{t+1} - \hat{\xi}_{t+1|t}). \tag{102}$$

We proceed as follows; first the Kalman filter in section 3.3.1 is calculated and sequences $\{\hat{\xi}_{t|t}\}_{t=1}^{T}, \{\hat{\xi}_{t+1|t}\}_{t=0}^{T-1}, \{\mathbf{P}_{t|t}\}_{t=1}^{T}$ and $\{\mathbf{P}_{t+1|t}\}_{t=1}^{T}$ are stored. Then (101) is used to generate $\{\mathbf{J}_{t|t}\}_{t=1}^{T-1}$. From this (102) is used to for $t = T - 1$ to calculate

$$\hat{\xi}_{T-1|T} = \hat{\xi}_{T-1|T-1} + \mathbf{J}_{T-1}(\hat{\xi}_{T|T} - \hat{\xi}_{T|T-1}).$$

Proceeding the same way for $T - 2$, the whole sequence of smoothed estimates $\{\hat{\xi}_{t|T}\}_{t=1}^{T}$ is calculated.

3.3.3 The maximum likelihood estimates

The forecasts $\hat{\xi}_{t+1|t}$ and $\hat{y}_{t+1|t}$ are optimal within the set of forecasts that are linear in $(x_t, \mathcal{Y}_{t-1})$. If the initial state ξ_1 and the innovations $\{v_t, w_t\}$ are multivariate Gaussian, then these forecasts are optimal among any functions of $(x_t, \mathcal{Y}_{t-1})$. Moreover, the conditional distribution of y_t on $(x_t, \mathcal{Y}_{t-1})$ is:

$$y_t|x_t, \mathcal{Y}_{t-1} \sim \mathrm{N}(\mathrm{A}x_t + \mathbf{H}\hat{\xi}_{t|t-1}, \mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H} + \mathrm{R}), \tag{103}$$

that is

$$\begin{aligned} &f_{y_t|x_t,\mathcal{Y}_{t-1}}(y_t|x_t, \mathcal{Y}_{t-1}) \\ &\quad = (2\pi)^{-1/2}|(\mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H} + \mathrm{R}|^{-1/2} \\ &\quad \times \exp\{-\frac{1}{2}(y_t - \mathrm{A}x_t + \mathbf{H}\hat{\xi}_{t|t-1})^2(\mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H} + \mathrm{R})^{-1}\} \\ &\quad \text{for } t = 1, 2, \ldots, T. \end{aligned} \tag{104}$$

The maximum likelihood is constructed by:

$$\sum_{t=1}^{T} \log f_{y_t|x_t,\mathcal{Y}_{t-1}}(y_t|x_t,\mathcal{Y}_{t-1}).$$

This expression is then maximized numerically with respect to $\mathbf{F}, \mathrm{A}, \mathbf{H}, R$ and $\mathbf{Q}$: See Burmeister and Wall (1982) for an illustrative application. Statistical inference of these estimates will be important in assessing the fit of the model and will be discussed in Section 6.

4 A profitability measure for a mean reverting spread

Let $f(0), f(\delta_k), \ldots, f(\delta_n)$ be the limiting density of $p_t \equiv \log(\alpha) + \log(p_t^A) - \log(p_t^B)$ at values $0 \leq \delta_k \leq \ldots \delta_n$. The value of δ_k that maximizes the profit is given by

$$\delta_{max} = \arg\max_{\delta_k}\{\delta_k f(\delta_k)\},$$

where n is chosen is sufficiently large. The expected profit for one trade is given that a position is entered where the spread is zero is:

$$E(p_t|p_t \geq \delta_{max}) = \frac{\sum_{\delta_k \geq \delta_{max}} \delta_k f(\delta_k)}{\sum_{\delta_k \geq \delta_{max}} f(\delta_k)} \tag{105}$$

.

5 Applying the search procedure

5.1 Treating historical and new information in the search procedure

At a given date the trader possesses practically infinite historical information on potential spreads on which the search procedure can be applied. However the most interesting spreads are those who show mean reversion up to the current date and therefore possibly would continue to show mean reversion in the future. The search procedure is initiated with 40 days of historical data. If the statistical tests for mean reversion are significant, the spread is extended with 20 days. This step is repeated until there is no significant evidence for mean reversion. The resulting spreads are then stored. When new information for these spreads becomes available, the tests are performed again to verify the mean reversion. Typically the trader would like to follow a self chosen set of spreads by excluding some spreads that are for some reasons not reliable. This subset is stored separately as Favorites in the search application.

5.2 Search criteria and results for the co-integration method

We must verify the existence of one stationary cointegrating relation described in the introduction to Section 2. The assumption of non-stationarity for the individual price series is verified if a unit-root test is not rejected. See Dickey, Said (1984) for a reference on unit root tests. The test described in section 2.4 is used to verify the hypothesis of zero cointegration relations against the alternative of 2. If this hypothesis is rejected at a 0.05 significance level, we must verify the hypothesis of 1 cointegration relation against the alternative of 2 relations. If this hypothesis cannot be rejected, we have evidence of one cointegration relation. Furthermore we must assess the second characteristic in Definition 3, Section 2, and thus reject a unit root in the series formed by the cointegrating relation. This is merely a verification of the cointegrating theory. Some of the results are found in Table 1. below, and the complete result table is found in appendix A.

5.3 Search criteria and results for the stochastic spread method

Besides information on the value of the spread, this method uses information from an industry index acting as a risk factor in the APT model, and the short (3 months) eurolib interest rate acting as the risk free rate. For this model to be evident, the OU model represented by the state equation (79) should show significance it its parameters, especially $F_{1,2}$. This is also the AR-parameter of the equation (79). If this parameter has a value very close to 0, the stationarity for the accumulated series is questionable, and therefore the hypothesis of a unit root has to be rejected at a significance level of 0.05 to verify the claim of stationarity. The table below show some of the pairs found under these criteria.

6 The software user interface

The images below show the user interface of the software application. It has three elementary functions:

i. Get new pairs - all available historical data is searched for potential pairs.

ii. Update Favorites - this function updates the Favorites subset of pairs with new data

iii. Update pairs - this function updates all previously found pairs with new data

Spread name	AR.parameter	Length	Delta	R_T	γ
dte gy /man gy	0.621	81	0.03	0.07	-1.02
dbk gy /man gy	0.812	171	0.02	0.05	-0.58
alv gy /man gy	0.765	171	0.02	0.06	-0.81
fp fp /muv2 gy	1	81	0.20	0.60	-3.33
fp fp /con gy	0.931	81	0.08	0.22	1.91
stm fp /man gy	0.713	81	0.03	0.07	-0.45
gle fp /bmw gy	0.703	81	0.02	0.04	-0.48
ml fp /vow gy	0.72	81	0.02	0.04	-0.84
fte fp /man gy	0.639	81	0.03	0.06	-1.11
dx fp /bay gy	0.736	111	0.03	0.06	-1.94
bn fp /vow gy	0.706	111	0.01	0.04	-0.38
bn fp /man gy	0.854	141	0.04	0.11	-0.71
bn fp /alv gy	0.678	111	0.01	0.04	-0.71
bn fp /fte fp	0.735	81	0.02	0.06	-0.47
agf fp /man gy	0.663	81	0.03	0.08	-1.02
wlsnc na /man gy	0.774	171	0.02	0.05	-0.71
wlsnc na /bn fp	0.632	111	0.02	0.06	-1.38
una na /man gy	0.655	81	0.01	0.03	-0.45

Table 1: The table shows some of the spreads where all criteria are met for the cointegration method. The first column shows the name of the spread with a three letter asset code and a two letter stock exchange code. The AR-parameter is shown verifying the stationarity of the spread. Its p-value of the standard t-test is given in parentheses, with 0 meaning negligible. The column Length shows the length in days for the spread. The column Delta shows the optimal trading point and R shows the expected profit of the spread in percentages of the exposed amount.

Spread name	AR.parameter	Length	Delta	R	ω
con gy /man gy	-0.086(0.067)	229	0.03	0.08	0.11(0.38)
con gy /dcx gy	0.867(0)	199	0.02	0.05	0.004(0.957)
bmw gy /dcx gy	-0.958(0)	259	0.00	0.00	-0.221(0)
rno fp /man gy	-0.073(0.098)	139	0.04	0.10	0.196(0.168)
rno fp /dcx gy	0.941(0)	169	0.02	0.05	0.01(0.874)
rno fp /con gy	-0.66(0.039)	79	0.01	0.01	0.286(0.066)
ug fp /vow gy	0.753(0.01)	349	0.06	0.24	-0.243(0)
ug fp /man gy	-1(0)	79	0.00	0.00	0.694(0.048)
ug fp /dcx gy	0.846(0.017)	259	0.05	0.14	-0.241(0.002)
ug fp /con gy	-0.992(0)	259	0.00	0.00	-0.15(0.138)
ug fp /bmw gy	-0.984(0)	199	0.00	0.00	-0.114(0.188)
ug fp /rno fp	-0.985(0)	229	0.00	0.00	-0.161(0.043)
ml fp /vow gy	-0.992(0)	379	0.00	0.00	-0.178(0.002)
ml fp /man gy	-0.935(0)	139	0.00	0.01	0.145(0.4)
ml fp /dcx gy	-0.624(0.073)	169	0.00	0.01	-0.009(0.897)
ml fp /bmw gy	-0.893(0)	379	0.00	0.00	0.076(0.137)
ml fp /rno fp	-0.583(0.052)	229	0.01	0.02	0.007(0.918)
ml fp /ug fp	-0.695(0.075)	319	0.00	0.01	0.061(0.397)

Table 2: The table shows spread where all criteria are met for the stochastic spread method. The columns in this table are identical to the columns in Table 2, except from ω that represent the weight of the risk factor. The AR-parameter now reflects the stationarity of the returns of the spread in equation (79).

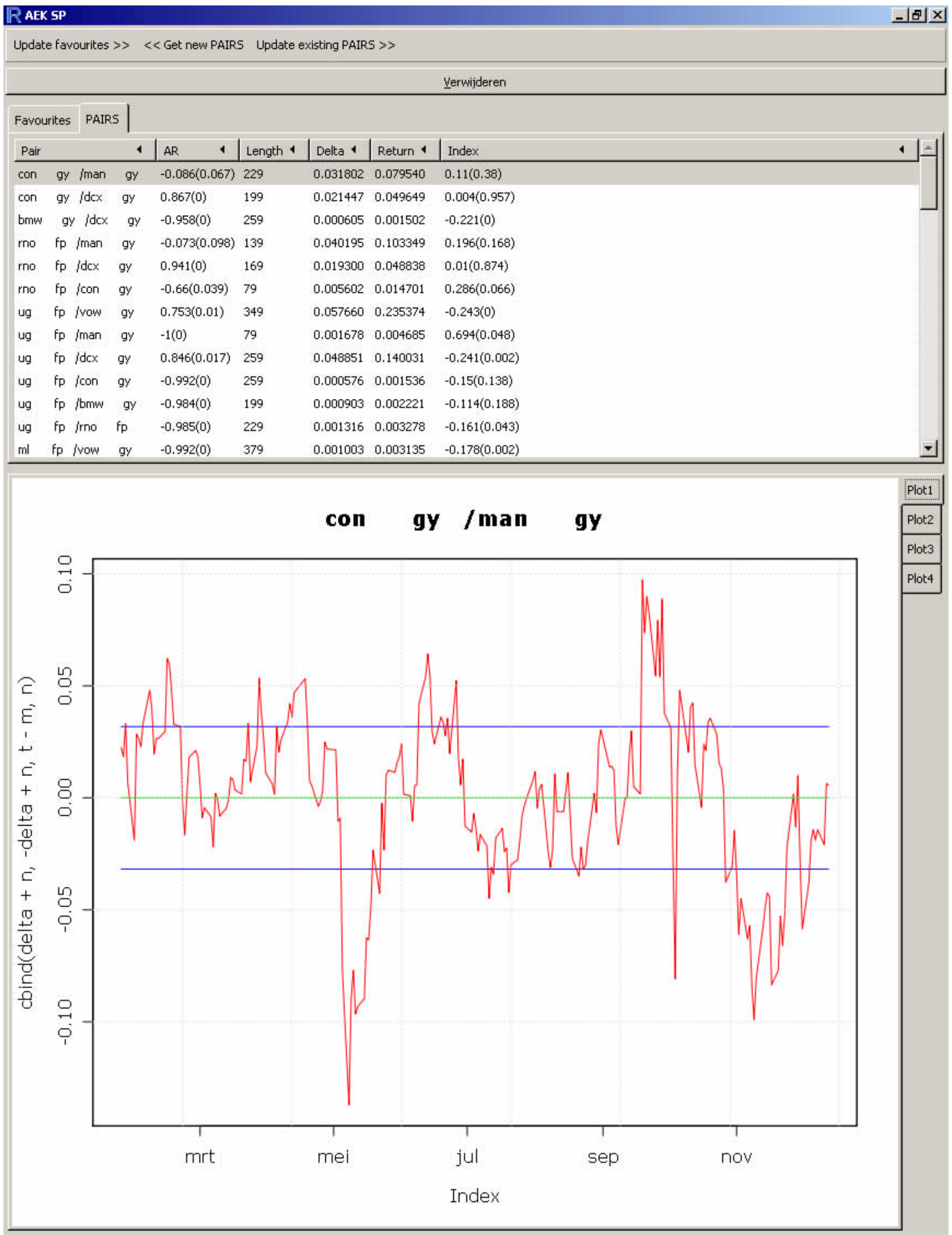

Figure 1: The user interface when using the stochastic spread method.

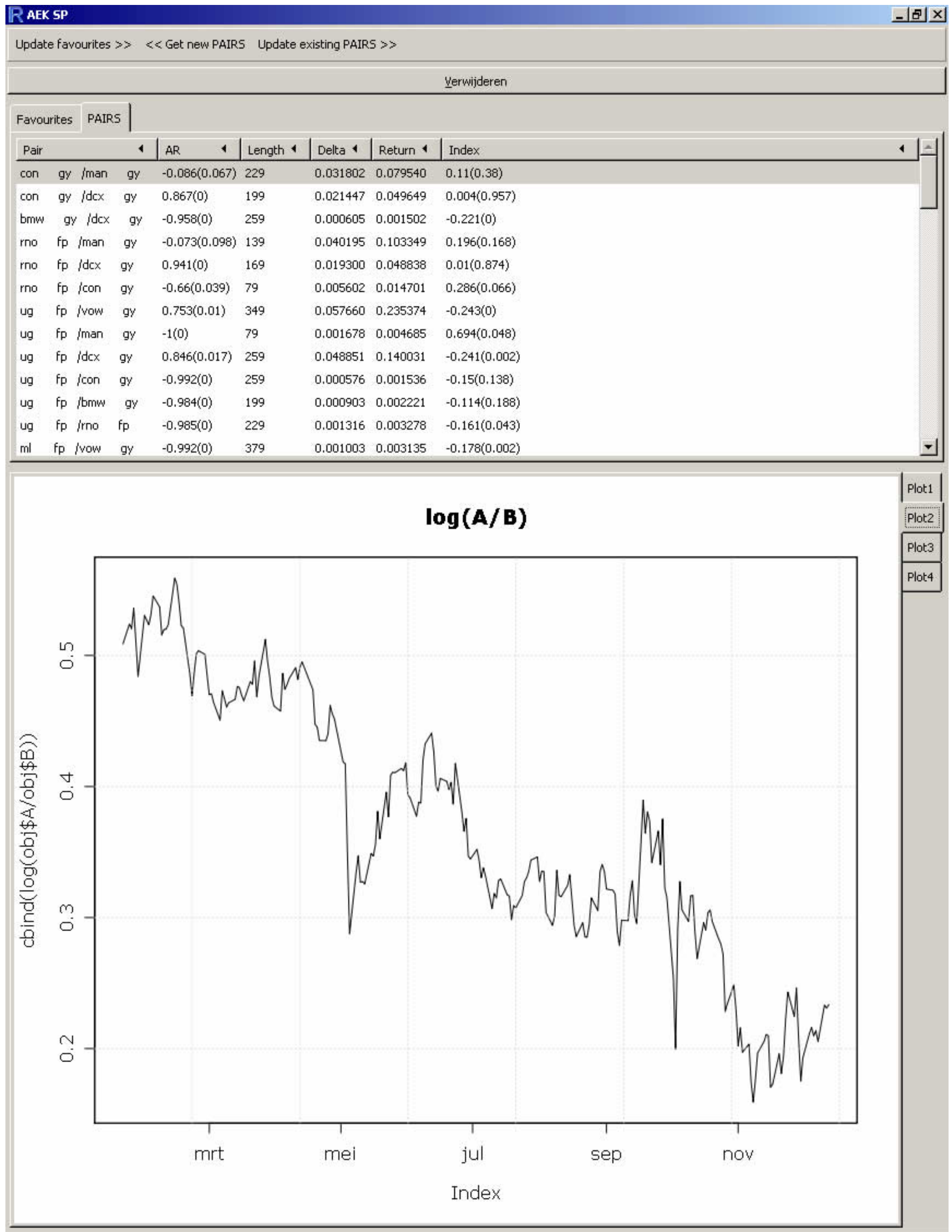

Figure 2: The user interface when using the stochastic spread method and selecting another plot view.

7 References

Do B., Faff R. and Hamza K. 2006. "A New Approach to Modeling and Estimation for Pairs Trading", Monash University, Working Paper.

Elliott, R., van der Hoek, J. and Malcolm, W. 2005. *"Pairs Trading"*, Quantitative Finance, Vol. 5(3), pp. 271-276.

Engle, R. and Granger, C. 1987. "Co-integration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, Vol. 55(2), pp. 251-276.

Hamilton, J. D. 1994. *"Time Series Analysis"*. Princeton: Princeton University Press.

Johansen, S. 1984, *"Functional Relations, Random Coefficients and Non-Linear Regression with Applications to Kinetic Data"*. Springer Verlag (125 pp).

Johansen, S. 1988, "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control* (v. 12, 1988), pp. 231-54.

Johansen, S. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, Econometric Society, vol. 59(6), pages 1551-80, November.

Kalman, R.E. 1960. "A New Approach to Linear Filtering and Prediction Problems", *Journal of Basic Engineering*, Vol. 82, pp.35-45.

Phillips, P C B and Durlauf, S N, 1986. "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, Blackwell Publishing, vol. 53(4), pages 473-95, August.

Ross, S. 1976, "The arbitrage theory of capital pricing", *Journal of Economic Theory*, v13, p341.

Said, S.E. and Dickey, D. A. 1984, "Testing for unit roots in autoregressive-moving average models of unknown order." *Biometrika*.; 71 (3) : 599-607

Vidyamurthy, G. 2004. *Pairs Trading, Quantitative Methods and Analysis*, John Wiley & Sons, Canada.

8 Appendix A - further results

8.1 Cointegration

Pair	AR.parameter	Length	R	Delta	γ
dte gy /man gy	0.621	81	0.03	0.07	-1.02
dbk gy /man gy	0.812	171	0.02	0.05	-0.58
alv gy /man gy	0.765	171	0.02	0.06	-0.81
fp fp /muv2 gy	1	81	0.20	0.60	-3.33
fp fp /con gy	0.931	81	0.08	0.22	1.91

Pair	AR.parameter	Length	R	Delta	γ
stm fp /man gy	0.713	81	0.03	0.07	-0.45
gle fp /bmw gy	0.703	81	0.02	0.04	-0.48
ml fp /vow gy	0.72	81	0.02	0.04	-0.84
fte fp /man gy	0.639	81	0.03	0.06	-1.11
dx fp /bay gy	0.736	111	0.03	0.06	-1.94
bn fp /vow gy	0.706	111	0.01	0.04	-0.38
bn fp /man gy	0.854	141	0.04	0.11	-0.71
bn fp /alv gy	0.678	111	0.01	0.04	-0.71
bn fp /fte fp	0.735	81	0.02	0.06	-0.47
agf fp /man gy	0.663	81	0.03	0.08	-1.02
wlsnc na /man gy	0.774	171	0.02	0.05	-0.71
wlsnc na /bn fp	0.632	111	0.02	0.06	-1.38
una na /man gy	0.655	81	0.01	0.03	-0.45
una na /stm fp	0.72	81	0.02	0.05	-1.11
tnt na /alv gy	0.806	171	0.04	0.15	-0.61
tnt na /bn fp	0.767	111	0.02	0.05	-0.99
sbmo na /man gy	0.896	81	0.09	0.24	-1.33
sbmo na /bas gy	0.751	141	0.02	0.06	-1.36
rdsa na /muv2 gy	0.978	81	0.16	0.43	-2.50
rdsa na /eoa gy	0.948	81	0.10	0.30	2.26
rdsa na /con gy	0.907	81	0.05	0.12	1.50
rdsa na /cbk gy	0.937	81	0.02	0.05	-1.19
rdsa na /gle fp	0.911	81	0.04	0.10	-1.84
phia na /stm fp	0.954	111	0.02	0.08	-0.60
phia na /una na	0.95	111	0.02	0.05	-0.84
kpn na /man gy	0.769	141	0.02	0.06	-0.67
kpn na /dte gy	0.787	81	0.01	0.04	-0.52
kpn na /alv gy	0.68	111	0.02	0.06	-0.84
kpn na /bn fp	0.607	111	0.02	0.05	-1.17
kpn na /wlsnc na	0.775	141	0.02	0.04	-0.95
heia na /stm fp	0.836	81	0.01	0.02	-0.35
heia na /rno fp	0.939	111	0.02	0.05	-0.15
heia na /cs fp	0.956	111	0.02	0.04	-0.05
heia na /una na	0.87	81	0.02	0.06	-0.08
gtn na /cbk gy	0.724	81	0.03	0.08	-2.11

Pair	AR.parameter	Length	R	Delta	γ
fora na /gle fp	0.857	111	0.05	0.17	-1.66
fora na /rdsa na	0.995	81	0.05	0.13	-0.40
dsm na /man gy	0.555	81	0.02	0.06	-0.98
dsm na /dbk gy	0.546	81	0.01	0.04	-1.64
dsm na /bay gy	0.807	81	0.04	0.11	-3.29
dsm na /stm fp	0.737	81	0.04	0.11	-2.37
dsm na /kpn na	0.717	81	0.04	0.11	-1.99
buhr na /su fp	0.817	81	0.07	0.22	-1.77
buhr na /ca fp	0.862	111	0.06	0.16	-3.83
asml na /ifx gy	0.915	81	0.11	0.28	-3.43
asml na /stm fp	0.826	141	0.02	0.05	-1.47
asml na /una na	0.852	81	0.02	0.05	-0.96
akza na /sap gy	0.925	81	0.12	0.46	1.72
akza na /bmw gy	0.894	81	0.06	0.19	0.64
akza na /gle fp	0.871	81	0.05	0.13	2.20
akza na /dx fp	0.909	81	0.06	0.17	0.82
akza na /cs fp	0.912	81	0.05	0.14	2.41
akza na /heia na	0.847	81	0.10	0.27	4.67
akza na /fora na	0.916	81	0.06	0.18	0.15
ah na /eoa gy	0.811	81	0.09	0.60	1.87
agn na /ah na	0.711	141	0.02	0.06	-0.58
aaba na /bay gy	0.709	81	0.02	0.05	-1.26

8.2 Stochastic spread

Spread name	AR.parameter	Length	Delta	R	ω
con gy /man gy	-0.086(0.067)	229	0.03	0.08	0.11(0.38)
con gy /dcx gy	0.867(0)	199	0.02	0.05	0.004(0.957)
bmw gy /dcx gy	-0.958(0)	259	0.00	0.00	-0.221(0)
rno fp /man gy	-0.073(0.098)	139	0.04	0.10	0.196(0.168)
rno fp /dcx gy	0.941(0)	169	0.02	0.05	0.01(0.874)
rno fp /con gy	-0.66(0.039)	79	0.01	0.01	0.286(0.066)
ug fp /vow gy	0.753(0.01)	349	0.06	0.24	-0.243(0)
ug fp /man gy	-1(0)	79	0.00	0.00	0.694(0.048)
ug fp /dcx gy	0.846(0.017)	259	0.05	0.14	-0.241(0.002)
ug fp /con gy	-0.992(0)	259	0.00	0.00	-0.15(0.138)
ug fp /bmw gy	-0.984(0)	199	0.00	0.00	-0.114(0.188)

Spread name	AR.parameter	Length	Delta	R	ω
ug fp /rno fp	-0.985(0)	229	0.00	0.00	-0.161(0.043)
ml fp /vow gy	-0.992(0)	379	0.00	0.00	-0.178(0.002)
ml fp /man gy	-0.935(0)	139	0.00	0.01	0.145(0.4)
ml fp /dcx gy	-0.624(0.073)	169	0.00	0.01	-0.009(0.897)
ml fp /bmw gy	-0.893(0)	379	0.00	0.00	0.076(0.137)
ml fp /rno fp	-0.583(0.052)	229	0.01	0.02	0.007(0.918)
ml fp /ug fp	-0.695(0.075)	319	0.00	0.01	0.061(0.397)
vie fp /rwe gy	0.675(0.092)	79	0.01	0.03	-0.16(0.378)
rdsa na /sbmo na	-0.096(0.026)	319	0.05	0.12	-0.295(0)
ren na /wlsnc na	-0.197(0.073)	79	0.03	0.08	-0.014(0.96)
fte fp /dte gy	-1(0)	79	0.00	0.00	-0.201(0.035)
kpn na /dte gy	-1(0)	79	0.00	0.00	-0.643(0)
buhr na /tnt na	0.932(0.001)	79	0.03	0.09	0.556(0.054)
buhr na /hgm na	0.885(0)	79	0.04	0.11	-0.282(0.451)
tms fp /sie gy	-0.122(0.049)	199	0.11	0.34	-0.082(0.274)
tms fp /sap gy	0.861(0.001)	289	0.07	0.22	-0.243(0.001)
tms fp /ifx gy	-0.76(0.034)	139	0.00	0.00	-0.286(0.01)
stm fp /tms fp	-0.788(0.099)	169	0.00	0.00	0.25(0.005)
su fp /tms fp	0.917(0)	349	0.03	0.09	-0.017(0.852)
cap fp /sap gy	-0.994(0)	319	0.00	0.00	0.143(0.023)
phia na /sie gy	0.893(0)	109	0.02	0.06	0.119(0.166)
phia na /dpw gy	0.815(0)	109	0.05	0.14	0.285(0.001)
phia na /tms fp	-0.639(0.001)	139	0.01	0.02	0.129(0.144)
phia na /stm fp	-0.999(0)	79	0.00	0.00	-0.512(0.001)
gtn na /su fp	0.68(0.001)	379	0.07	0.20	-0.122(0.365)
asml na /sap gy	-0.948(0)	349	0.00	0.00	0.026(0.638)
asml na /tms fp	-0.867(0.002)	199	0.00	0.00	0.255(0.002)
asml na /su fp	-0.984(0)	319	0.00	0.00	0.251(0.002)
ai fp /bay gy	-0.716(0.047)	319	0.00	0.00	-0.363(0)
dsm na /bay gy	0.9(0)	379	0.03	0.09	-0.482(0)
dsm na /ai fp	0.913(0)	109	0.04	0.10	-0.297(0.256)
akza na /bay gy	-0.614(0.065)	169	0.00	0.01	-0.094(0.377)
akza na /bas gy	-0.646(0.069)	319	0.00	0.00	-0.02(0.812)
akza na /dsm na	0.866(0)	319	0.05	0.23	0.147(0.114)
bn fp /meo gy	-0.884(0)	289	0.00	0.00	-0.109(0.182)

Spread name	AR.parameter	Length	Delta	R	ω
una na /bn fp	-0.852(0)	319	0.00	0.00	-0.102(0.154)
num na /meo gy	-0.83(0.013)	79	0.00	0.00	-0.486(0.119)
num na /una na	-0.116(0.006)	289	0.04	0.09	-0.077(0.44)
heia na /una na	-0.877(0)	79	0.00	0.00	-0.34(0.012)
heia na /num na	-0.91(0)	349	0.00	0.00	-0.165(0.123)
ah na /bn fp	-0.686(0)	379	0.00	0.01	0.432(0)
ah na /num na	-0.701(0.089)	319	0.00	0.01	0.576(0)
cs fp /dx fp	0.744(0.034)	199	0.02	0.05	0.602(0)
agf fp /muv2 gy	-1(0)	109	0.00	0.00	-0.265(0.009)
inga na /agf fp	-1(0)	79	0.00	0.00	0.576(0)
fora na /agf fp	-0.992(0)	139	0.00	0.00	0.198(0.025)
agn na /gle fp	0.108(0.04)	199	0.04	0.12	0.158(0.041)
agn na /cs fp	0.765(0.005)	199	0.03	0.07	-0.038(0.594)
cbk gy /dbk gy	-0.996(0)	379	0.00	0.00	0.158(0.068)
aca fp /cbk gy	-0.981(0)	349	0.00	0.00	-0.168(0.104)
bnp fp /dbk gy	0.742(0.042)	169	0.04	0.14	0.009(0.878)
inga na /dbk gy	-0.992(0)	79	0.00	0.00	0.078(0.662)
inga na /cbk gy	-1(0)	349	0.00	0.00	-0.313(0.003)
fora na /cbk gy	-0.978(0)	289	0.00	0.01	-0.324(0.003)
aaba na /cbk gy	-0.979(0)	259	0.00	0.01	-0.51(0)
aaba na /dx fp	0.659(0.01)	229	0.03	0.06	0.121(0.049)