2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Optimal Mean-Variance Portfolio Construction in Cointegrated Vector Autoregressive Systems

Melanie B. Rudoy, Charles E. Rohrs
Massachusetts Institute of Technology
Digital Signal Processing Group
77 Massachusetts Avenue, Cambridge, MA 02139
{mbs, crohrs}@mit.edu

Abstract— **We study the problem of optimal portfolio construction when the log-prices follow a discrete-time cointegrated vector autoregressive model. We follow the classical Markowitz mean-variance optimization approach, and derive expressions for the optimal portfolio weight vector over a single decision interval, both for a finite-time horizon and in the limit of an infinite horizon. It is often stated in the literature that given assets whose price dynamics exhibit cointegration, the portfolio weights should be chosen from the space of cointegrating relations, resulting in what is commonly referred to as the *beta* portfolio. However, we show here that the optimal action in the mean-variance sense for a finite trading interval is to buy the portfolio with a component both in the beta direction and a component in the direction of expected change. Furthermore, we prove that the beta portfolio is optimal only in the limit of an infinite trading horizon. Additionally, we derive the conditions under which the optimal portfolio is proportional to the disequilibrium readjustment forces of the cointegration model. Our results rely on a careful eigenanalysis of the underlying state space model, in which we derive a closed form solution for the optimal Markowitz portfolio, which is well-behaved despite the nonstationarity of the underlying price dynamics. We demonstrate our results with evaluations using both simulated and historical data.**

I. Introduction

Over the last three decades, many authors have shown that there exist groups of real-world economic time series that follow a vector autoregressive (VAR) process, and that these signals may share one or more unit roots, a property known as cointegration [1], [2]. While each of the underlying signals of the vector process is nonstationary due to the random walk component, the corresponding first difference series are wide-sense stationary. Furthermore, when a VAR model exhibits cointegration, it is possible to construct a linear combination of the underlying time series that is stationary, by choosing coefficients from within the space of cointegrating vectors. In this paper, it is assumed that a set of cointegrated financial products has been identified through some means, such as the methods described in [3], and we address the question of how to construct portfolios using only this universe of assets.

The trading of cointegrated assets has been previously discussed in the literature [4]–[6]. A common theme within these works is the reliance on statistical arbitrage techniques for trading the stationary linear combination, such as the methods described in [4], [7]. One such technique is a mean-reverting scheme, in which the entire portfolio is bought when the stationary signal deviates from its mean by a predetermined threshold, and the position is closed when the signal mean reverts. Here, we show that portfolios bought purely in the direction of a cointegrating vector are not optimal in the traditional Markowitz mean-variance sense for single-period, finite trading horizons, and we derive a closed-form expression for the optimal portfolio.

The asset allocation rule derived here maximizes the expected return on the portfolio given a constraint on the variance of the return, for a fixed time horizon, under the assumption that rebalancing at intermediate times is prohibited. It is commonly believed that constructing mean-variance optimal portfolios in cointegrated systems is ill-posed due to the fact that the underlying dynamics are nonstationary. In particular, the random variable corresponding to the change in the log-prices has a covariance matrix that diverges as a function of the trading horizon. However, we show that there is an additional, positive expected return to be gained from choosing the portfolio not only in the direction of finite variance, but also in the direction of expected change. We show that only in the limit of an infinite trading horizon, do the portfolio weights asymptotically approach those in the cointegrating space. It is also shown that under a slightly modified set of assumptions, the optimal portfolio weights are proportional to the disequilibrium readjustment forces of the cointegration model. Our results are consistent with the continuous-time solution given in [8].

The organization of this paper is as follows. Section II contains an overview of cointegrated vector autoregressive models, their representation in state-space form, and classical Markowitz portfolio theory. In Section III, a closed-form expression for the mean-variance optimal portfolio is derived as a function of the trading horizon and the solution in the limit of an infinite horizon is presented. In Section IV, the asset allocation rule for the case where the variance constraint is replaced by a leverage constraint is given. The analysis of a synthetic example is discussed in Section V, and finally the results of a trading simulation using real, historical data are presented in Section VI.

978-1-4244-2079-7/08/$25.00 ©2008 AACC.

II. Problem Formulation

Let $\mathbf{x}_i$ be a p-dimensional random vector representing the log-prices of a set of assets, that obey the following k^{th} order vector autoregressive process:

$$\mathbf{x}_i = \mathbf{\Pi}_1 \mathbf{x}_{i-1} + \ldots + \mathbf{\Pi}_k \mathbf{x}_{i-k} + \mathbf{\Phi}\mathbf{d}_i + \boldsymbol{\epsilon}_i. \tag{1}$$

Here the $p \times p$ $\mathbf{\Pi}_j$, $j \in 1 \ldots k$ matrices relate the current value of each component process to the lagged versions of the other processes, $\mathbf{d}_i$ is an r-dimensional vector of deterministic inputs, $\mathbf{\Phi}$ is a $p \times r$ matrix of coefficients relating the deterministic inputs to the elements of $\mathbf{x}_i$, and $\boldsymbol{\epsilon}_i$ is a p-dimensional Gaussian random noise vector with zero mean and variance $\mathbf{\Psi}$. We refer to a model of this form with k lagged terms as a VAR(k) model. In the general VAR framework, it is possible to specify constraints on the matrices $\mathbf{\Pi}_j$ so that each component time series is wide sense stationary. It is also possible to specify conditions so that the overall system exhibits a special form of nonstationarity, known as cointegration. This occurs when the matrix $\mathbf{\Pi} \triangleq \sum_{j=1}^{k} \mathbf{\Pi}_j - \mathbf{I}$ is not full rank, due to the presence of at least one pole on the unit circle, known as a unit root. Throughout this paper, we assume the unit roots are located at $z = 1$. We can express $\mathbf{\Pi}$ as the outer product of two $p \times r$ matrices, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, with $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^T$, where $r < p$ denotes the reduced rank of $\mathbf{\Pi}$. The column space of $\boldsymbol{\beta}$ is commonly referred to as the cointegrating space, and the vectors in the column space of $\boldsymbol{\alpha}$ are referred to as the disequilibrium adjustment forces. As a result of the common unit roots, each component of $\mathbf{x}_i$ is nonstationary, but it can be shown that for all $\mathbf{b}$ in the span of $\{\boldsymbol{\beta}\}$, $\mathbf{b}^T\mathbf{x}_i$ is wide-sense stationary [2].

Equation (1) may equivalently be expressed in state-space form by augmenting the state vector with all of the lagged terms of the process. When $p = k = 2$, we have:

$$\underbrace{\begin{bmatrix} \mathbf{x}[i] \\ \mathbf{x}[i-1] \end{bmatrix}}_{\mathbf{q}[i+1]} = \underbrace{\begin{bmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 \\ \mathbf{I}_2 & \mathbf{0} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{x}[i-1] \\ \mathbf{x}[i-2] \end{bmatrix}}_{\mathbf{q}[i]} + \underbrace{\begin{bmatrix} \mathbf{I}_2 \\ \mathbf{0} \end{bmatrix}}_{\mathbf{b}} \underbrace{\begin{bmatrix} \mathbf{\Phi}\mathbf{d}_i + \boldsymbol{\epsilon}_i \end{bmatrix}}_{\mathbf{u}[i]} \tag{2}$$

$$\underbrace{\mathbf{x}[i]}_{y[i]} = \underbrace{\begin{bmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 \end{bmatrix}}_{\mathbf{c}^T} \underbrace{\begin{bmatrix} \mathbf{x}[i-1] \\ \mathbf{x}[i-2] \end{bmatrix}}_{\mathbf{q}[i]} + \underbrace{\begin{bmatrix} \mathbf{I}_2 \end{bmatrix}}_{\mathbf{d}} \underbrace{\begin{bmatrix} \mathbf{\Phi}\mathbf{d}_i + \boldsymbol{\epsilon}_i \end{bmatrix}}_{\mathbf{u}[i]},$$

where $\mathbf{I}_k$ is the k dimensional identity matrix. Throughout this paper, we rely extensively on the Jordan canonical form of the state-space model, where the modes are decoupled as much as possible. The state transition matrix $\mathbf{A}$ is factored as $\mathbf{A} = \mathbf{M}\mathbf{J}\mathbf{M}^{-1}$, where $\mathbf{J}$ is a matrix of Jordan blocks containing the eigenvalues of the system and $\mathbf{M}$ is a matrix of column vectors containing the generalized eigenvectors, $\mathbf{m}_i$, which are linearly independent, but not necessarily orthogonal [9]. By construction, the matrix $\mathbf{A}$ is $n \times n$, where $n = pk$. We assume that each underlying series is nonstationary, but that the corresponding first difference series are wide-sense stationary, and thus all of the eigenvalues must lie either inside the unit circle or at $z = 1$. Specifically, let us assume that $\lambda_1 = 1$ and $|\lambda_k| < 1$ for all $k = 2 \ldots n$. As a consequence of the special block matrix structure for $\mathbf{A}$ given in Eq. (2), and the additional assumption that the geometric and algebraic multiplicities for each eigenvalue coincide, the n eigenvectors have the following block form:

$$\mathbf{m}_i = \begin{pmatrix} \widetilde{\mathbf{m}}_i^T & \lambda_i^{-1}\widetilde{\mathbf{m}}_i^T & \ldots & \lambda_i^{-(k-1)}\widetilde{\mathbf{m}}_i^T \end{pmatrix}^T, \tag{3}$$

where each $\widetilde{\mathbf{m}}_i$ is a $p \times 1$ vector, as described in [2]. Subsequently, we shall refer to the $\widetilde{\mathbf{m}}_i$ as the *base vectors* of the eigenvectors of $\mathbf{A}$.

In order to determine the portfolio weights when the log-prices for the underlying assets follow a cointegrated VAR model, we adopt the classical Markowitz portfolio optimization approach [10]. Let $\mathbf{x}_i$ be a random vector representing the current log prices of p assets at time step i, where the initial log price is given by the vector $\mathbf{x}_0$. Also let $\Delta\mathbf{x} = \mathbf{x}_N - \mathbf{x}_0$ denote the random variable representing the change in log price of each asset over a single decision period, corresponding to a fixed trading horizon of length N. Since $\mathbf{x}_i$ follows the Gaussian random process given in Eq. (1), $\Delta\mathbf{x}$ is also Gaussian with mean $\boldsymbol{\mu}_N$ and covariance matrix $\mathbf{C}_N$. We maximize the expected portfolio return for a trading horizon of length N, given an upper bound on the allowable portfolio risk, using the following quadratic program, P_0:

$$\left.\begin{array}{ll} \mathbf{w}^* = \arg\max_{\mathbf{w}} & \mathbf{w}^T\boldsymbol{\mu}_N \\ \text{subject to} & \mathbf{w}^T\mathbf{C}_N\mathbf{w} \leq \sigma_0^2, \end{array}\right\} \; P_0 \tag{4}$$

where the portfolio weight vector, $\mathbf{w}$, denotes the percentage of initial wealth to allocate to each asset. A weight with a positive sign denotes a long position, while a weight with a negative sign denotes a short position. We allow the overall portfolio to be leveraged, i.e. the market value of the portfolio at the entry point may exceed the initial wealth available, and therefore a constraint of the form $\mathbf{1}^T\mathbf{w} = 1$ is not required. The degree to which the portfolio is leveraged is limited by the allowable risk parameter, σ_0. It is well known that the solution to problem P_0 is given by

$$\mathbf{w}^* = \frac{\sigma_0}{\sqrt{\boldsymbol{\mu}_N^T\mathbf{C}_N^{-1}\boldsymbol{\mu}_N}}\mathbf{C}_N^{-1}\boldsymbol{\mu}_N. \tag{5}$$

In the next section, we derive closed-form expressions for both $\mathbf{C}_N^{-1}$ and $\boldsymbol{\mu}_N$ for a cointegrated VAR system, and show that $\mathbf{w}^* = \boldsymbol{\beta}$ only in the limit as N approaches infinity.

III. Optimal Portfolio Construction

We consider a universe of p financial assets whose underlying log-prices follow a cointegrated VAR(k) model, with no exogeneous stochastic inputs and a constant deterministic input, so that $\mathbf{\Phi}$ is a $p \times 1$ vector and $\mathbf{d}_i = 1$. The constant terms are included in the model in order to capture the overall

linear growth trend present in the historical log-prices of most assets. The mean and variance of the change in log price of the assets over a period of length N is given by:

$$E[\Delta \mathbf{x}] = \boldsymbol{\mu}_N = \mathbf{T}\left[\left(\mathbf{A}^N - \mathbf{I}\right)\mathbf{x}_0\right] + N\boldsymbol{\Phi} \quad (6)$$

$$\text{Var}[\Delta \mathbf{x}] = \mathbf{C}_N = \mathbf{T}\left[\sum_{i=0}^{N-1} \mathbf{A}^i \widetilde{\boldsymbol{\Psi}} \left(\mathbf{A}^i\right)^T\right]\mathbf{T}^T, \quad (7)$$

where $\mathbf{T} = \begin{pmatrix}\mathbf{I}_p & \mathbf{0}_{p\times n-p}\end{pmatrix}$ and $\widetilde{\boldsymbol{\Psi}}$ is a $pk \times pk$ matrix given by $\widetilde{\boldsymbol{\Psi}} = \text{diag}\left(\boldsymbol{\Psi}, \mathbf{0}_p, ..., \mathbf{0}_p\right)$.

We seek an expression for $\boldsymbol{\mu}_N$ as a function of the eigenvalues and eigenvectors of the overall system. Equation (6) can be rewritten as:

$$\begin{aligned}
\boldsymbol{\mu}_N &= \mathbf{TM}\left(\mathbf{J}^N - \mathbf{I}\right)\mathbf{M}^{-1}\mathbf{x}_0 + N\boldsymbol{\Phi} \\
&= \mathbf{T}\sum_{i=1}^{n=pk} c_i(\lambda_i^N - 1)\mathbf{m}_i + N\boldsymbol{\Phi} \\
&= \sum_{i=1}^{n} c_i(\lambda_i^N - 1)\mathbf{T}\mathbf{m}_i + N\boldsymbol{\Phi} \\
&= \sum_{i=2}^{n} c_i(\lambda_i^N - 1)\widetilde{\mathbf{m}}_i + N\boldsymbol{\Phi}, \quad (8)
\end{aligned}$$

where the c_i's are the expansion coefficients of $\mathbf{x}_0$ in the basis defined by $\{\mathbf{m}_1, ..., \mathbf{m}_n\}$ and the last step follows from Eq. (3) and the fact that $\lambda_1 = 1$. Thus we see that the direction of expected change can be expressed as a function of the base vectors used to describe the block structure of the eigenvectors of the state transition matrix $\mathbf{A}$ from Eq. 2.

We now turn our attention to understanding the behavior of covariance matrix of $\Delta\mathbf{x}$ as a function of the trading horizon, which can be computed using a matrix difference equation, as:

$$\begin{aligned}
\widetilde{\mathbf{C}}_N &= \mathbf{A}\widetilde{\mathbf{C}}_{N-1}\mathbf{A}^T + \widetilde{\boldsymbol{\Psi}} \quad (9)\\
\mathbf{C}_N &= \mathbf{T}\widetilde{\mathbf{C}}_N\mathbf{T}^T.
\end{aligned}$$

In order for Eq. (9) to have a steady-state solution, $\mathbf{C}$, it must satisfy the discrete-time Lyapunov equation, given by:

$$\mathbf{C} - \mathbf{ACA}^T - \widetilde{\boldsymbol{\Psi}} = 0$$

However, due to fact that $\mathbf{A}$ has an eigenvalue at unity, the difference equation is unstable and $\widetilde{\mathbf{C}}_N$ has one eigenvalue that diverges as N increases. Fortunately, the optimal portfolio weights do not directly depend on $\mathbf{C}_N$, but rather on $\mathbf{C}_N^{-1}$, which is well behaved. Theorem 3.1 describes the behavior of the eigenvectors and eigenvalues of both $\mathbf{C}_N$ and $\mathbf{C}_N^{-1}$ as a function of the trading horizon, N.

Theorem 3.1: When $N = 1$, the eigenvectors of $\mathbf{C}_N$ and $\mathbf{C}_N^{-1}$ are aligned with the eigenvectors of $\boldsymbol{\Psi}$. These eigenvectors converge to $\{\boldsymbol{\beta}, \boldsymbol{\beta}_\perp\}$ as N approaches infinity. The eigenvalue associated with $\boldsymbol{\beta}$ converges to a strictly positive, real-valued scalar, while the eigenvalue associated with $\boldsymbol{\beta}_\perp$ diverges in $\mathbf{C}_N$ and converges to zero in $\mathbf{C}_N^{-1}$.

Proof: Letting $\widetilde{\boldsymbol{\Psi}} = \mathbf{SS}^T$ and using Eq. (7), the covariance matrix for $\Delta\mathbf{x}$ after N periods is:

$$\begin{aligned}
\mathbf{C}_N &= \mathbf{T}\left[\sum_{i=0}^{N-1} \mathbf{A}^i \widetilde{\boldsymbol{\Psi}} \left(\mathbf{A}^i\right)^T\right]\mathbf{T}^T \\
&= \mathbf{T}\left[\sum_{i=0}^{N-1} \mathbf{MJ}^i\mathbf{M}^{-1}\mathbf{SS}^T\left(\mathbf{M}^{-1}\right)^T\mathbf{J}^i\mathbf{M}^T\right]\mathbf{T}^T \\
&= \mathbf{TM}\left[\sum_{i=0}^{N-1} \left(\mathbf{J}^i\mathbf{M}^{-1}\mathbf{S}\right)\left(\mathbf{J}^i\mathbf{M}^{-1}\mathbf{S}\right)^T\right]\mathbf{M}^T\mathbf{T}^T \\
&= \mathbf{TM}\sum_{i=0}^{N-1}\begin{pmatrix} c_{1,1}^i\mathbf{q}_1^T\mathbf{q}_1 & \cdots & c_{1,n}^i\mathbf{q}_1^T\mathbf{q}_n \\ c_{2,1}^i\mathbf{q}_2^T\mathbf{q}_1 & \cdots & c_{2,n}^i\mathbf{q}_2^T\mathbf{q}_n \\ \vdots & & \vdots \\ c_{n,1}^i\mathbf{q}_n^T\mathbf{q}_1 & \cdots & c_{n,n}^i\mathbf{q}_n^T\mathbf{q}_n \end{pmatrix}(\mathbf{TM})^T,
\end{aligned}$$

where $c_{i,j} = \lambda_i\lambda_j$ and $\mathbf{q}_i$ is the i^{th} column of the matrix $\mathbf{Q} = \mathbf{M}^{-1}\mathbf{S}$. Using the fact that $\lambda_1 = 1$, we can evaluate the summation as $\widetilde{\mathbf{C}}_N = \mathbf{MKM}^T$, where:

$$\mathbf{K} = \begin{pmatrix} N\mathbf{q}_1^T\mathbf{q}_1 & \cdots & \frac{1-c_{1,n}^N}{1-c_{1,n}}\mathbf{q}_1^T\mathbf{q}_n \\ \frac{1-c_{2,1}^N}{1-c_{2,1}}\mathbf{q}_2^T\mathbf{q}_1 & \cdots & \frac{1-c_{2,n}^N}{1-c_{2,n}}\mathbf{q}_2^T\mathbf{q}_n \\ \vdots & & \vdots \\ \frac{1-c_{n,1}^N}{1-c_{n,1}}\mathbf{q}_n^T\mathbf{q}_1 & \cdots & \frac{1-c_{n,n}^N}{1-c_{n,n}}\mathbf{q}_n^T\mathbf{q}_n \end{pmatrix}.$$

Multiplying through we get:

$$\mathbf{C}_N = \sum_{i=1}^{n}\sum_{j=1}^{n} K_{i,j}\widetilde{\mathbf{m}}_i\widetilde{\mathbf{m}}_j^T.$$

As N approaches infinity, the first term, $N\mathbf{q}_1^T\mathbf{q}_1\widetilde{\mathbf{m}}_1\widetilde{\mathbf{m}}_1^T$, dominates the summation, causing the covariance matrix to diverge in the direction of $\widetilde{\mathbf{m}}_1 = \boldsymbol{\beta}_\perp$. Hence $\mathbf{C}_\infty$ has one eigenvector in the direction of $\boldsymbol{\beta}_\perp$ with corresponding eigenvalue of infinity, and the second eigenvector in the direction of $\boldsymbol{\beta}$ with a bounded eigenvalue denoted by γ. The inverse covariance matrix $\mathbf{C}_\infty^{-1}$ has the same eigenvectors as $\mathbf{C}_N$, with eigenvalues of zero and $\frac{1}{\gamma}$, respectively. ■

As a result of the zero eigenvalue in the direction of $\boldsymbol{\beta}_\perp$, the optimal portfolio in the limit of an infinite trading horizon is in the direction of $\boldsymbol{\beta}$, independent of the direction of $\boldsymbol{\mu}_\infty$.

IV. Leverage Constraint

In this section, we address how to construct the portfolio when the variance constraint is replaced by a leverage constraint, i.e. a constraint on the length of the portfolio vector, such as $\mathbf{w}^T\mathbf{w} = 1$. We find that the optimal action is to choose the portfolio in the direction of expected change, which in certain cases may be equal to the $\boldsymbol{\alpha}$ vector.

Theorem 4.1: Given a constraint on the degree of portfolio leverage, the optimal portfolio weight vector for a trading horizon of length N is proportional to the direction of expected change, as:

$$\mathbf{w}_N^* \propto \boldsymbol{\mu}_N = \sum_{i=2}^{n} c_i(\lambda_i^N - 1)\widetilde{\mathbf{m}}_i + N\boldsymbol{\Phi},$$

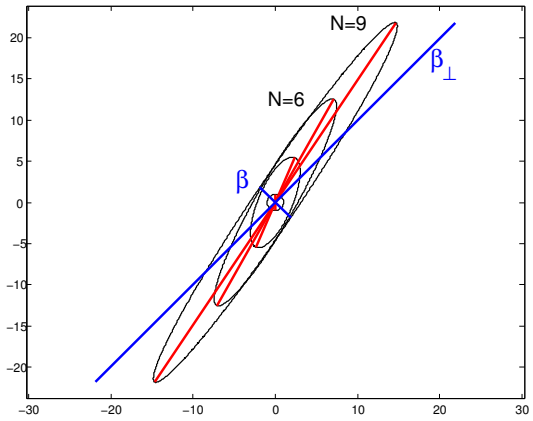


Fig. 1. Evolution of Covariance Matrix Principal Axes. The lengths of the axes are proportional to the corresponding eigenvalues.

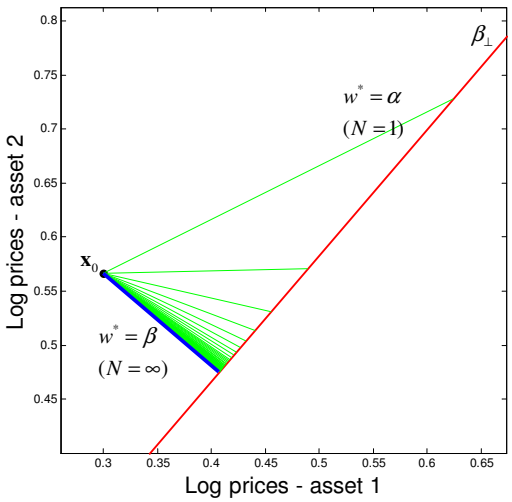


Fig. 2. Evolution of direction of optimal portfolio weight vectors. The weights start in the direction of $\boldsymbol{\alpha}$, and converge to the direction of $\boldsymbol{\beta}$ as the trading horizon (N) increases.

where the $\widetilde{\mathbf{m}}_i$ are the base vectors used to construct the block form of the generalized eigenvectors, as defined in Eq. (3). In the limit of an infinite trading horizon and when $\boldsymbol{\Phi} = \mathbf{0}^T$, the optimal weight vector is in the direction of

$$\mathbf{w}^*_\infty \propto -\sum_{i=2}^{n} c_i \widetilde{\mathbf{m}}_i.$$

Proof: According to the Cauchy-Schwartz inequality, the inner product $\mathbf{w}^T \boldsymbol{\mu}_N$ subject to $\mathbf{w}^T\mathbf{w} = 1$ is maximized when the vector $\mathbf{w}$ is chosen to be proportional to $\boldsymbol{\mu}_N$. By Eq. (8) and the fact that $|\lambda_i| < 1$ for $i = 2 \ldots n$ and $\boldsymbol{\Phi} = \mathbf{0}^T$, the limit follows. ■

As a special case, we consider a two asset system following a cointegrated VAR(1) model. When $\boldsymbol{\Phi} = \mathbf{0}^T$, the direction of the optimal portfolio is equal to the disequillibrium readjustment force, $\boldsymbol{\alpha}$, as stated in Theorem 4.2.

Theorem 4.2: Given a cointegrated VAR system with $k = 1$ and $p = 2$ and $\boldsymbol{\Phi} = \mathbf{0}^T$, the optimal leverage-constrained portfolio is proportional to $\boldsymbol{\alpha}$ for all N.

Proof: Recall that the matrix $\boldsymbol{\Pi}$ can be factored as:

$$\begin{aligned}
\boldsymbol{\Pi} &= \boldsymbol{\Pi}_1 - \mathbf{I}_2 = \mathbf{M}\mathbf{J}\mathbf{M}^{-1} - \mathbf{I}_2 \\
&= \begin{pmatrix}\mathbf{m}_1 & \mathbf{m}_2\end{pmatrix} \begin{pmatrix}0 & 0 \\ 0 & \lambda - 1\end{pmatrix} \begin{pmatrix}\mathbf{m}_1 & \mathbf{m}_2\end{pmatrix}^{-1} \\
&= \begin{pmatrix}\mathbf{m}_1 & \mathbf{m}_2\end{pmatrix} \begin{pmatrix}0 & 0 \\ 0 & \lambda - 1\end{pmatrix} \begin{pmatrix}\mathbf{n}_1 \\ \mathbf{n}_2\end{pmatrix} \\
&= (\lambda - 1)\, \mathbf{m}_2 \mathbf{n}_2^T,
\end{aligned}$$

where $\mathbf{n}_k$ is the k^{th} row of $\mathbf{M}^{-1}$. We also know that $\boldsymbol{\Pi}$ has rank 1, and it therefore can be written as the outer product of two 2×1 vectors, as $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^T$. Equating both factorizations, we see that $\mathbf{m}_2$ must be proportional to $\boldsymbol{\alpha}$. According to Theorem 4.1, $\mathbf{w}^*_N$ will be proportional to $\mathbf{m}_2$ for all N, and hence proportional to $\boldsymbol{\alpha}$. ■

V. Simulation Results

Let us now consider a synthetic example for a VAR(1) system of two assets with input driving covariance given by $\boldsymbol{\Psi} = \mathbf{I}$, and no deterministic inputs, i.e. $\boldsymbol{\Phi} = \mathbf{0}^T$. The principal axes of $\mathbf{C}_N$ are initially aligned with the unit vectors in the plane, and converge to $\{\boldsymbol{\beta}, \boldsymbol{\beta}_\perp\}$ as N increases, as depicted in Figure 1. The eigenvalue associated with $\boldsymbol{\beta}$ converges to $\gamma = \frac{1}{1-\lambda^2}$, while the eigenvalue associated with $\boldsymbol{\beta}_\perp$ diverges. The inverse covariance matrix, $\mathbf{C}_N^{-1}$, has the same eigenvectors as $\mathbf{C}_N$, but eigenvalues that approach $1 - \lambda^2$ and zero. The mean-variance optimal portfolio weights for this example can be computed as a function of N using Eq. (5). We find that for $N = 1$ the weights are proportional to $\boldsymbol{\alpha}$, and converge to $\boldsymbol{\beta}$ as the trading horizon increases, as depicted in Figure 2. The initial log price pair was chosen to be $\mathbf{x}_0 = \begin{pmatrix}0.3 & 0.5\end{pmatrix}^T$, which represents a state of mispricing relative to the long-term equilibrium vector, $\boldsymbol{\beta}_\perp$. Only in the limit of an infinite horizon is the optimal portfolio in the direction of $\boldsymbol{\beta}$.

In Figure 3, we explore the mean-variance tradeoff of various portfolios by utilizing the concept of a leverage constraint, as discussed in Section IV. We compare the expected return as a function of trading horizon for the three portfolios corresponding to $\mathbf{w} = \boldsymbol{\alpha}$, $\mathbf{w} = \boldsymbol{\beta}$, and $\mathbf{w} = \mathbf{w}^*$, the mean-variance optimal portfolio, with each normalized so that $||\mathbf{w}||_2 = 1$. Again, the initial log price pair was chosen to be $\mathbf{x}_0 = \begin{pmatrix}0.3 & 0.5\end{pmatrix}^T$. The highest expected return is achieved with the $\boldsymbol{\alpha}$ portfolio, due to the fact that in a VAR(1) model with two assets $\boldsymbol{\mu}_N$ is proportional to $\boldsymbol{\alpha}$ independent of trading horizon, however, the variance of this portfolio grows linearly with increasing N. The $\boldsymbol{\beta}$ portfolio has smaller expected return, but the variance converges to a finite quantity. The optimal portfolio is aligned with $\boldsymbol{\alpha}$ for small N, but as N increases and the variance grows, the optimal weight vector $\mathbf{w}^*$ is pulled toward $\boldsymbol{\beta}$ in order to satisfy the variance constraint, until it is perfectly aligned with $\boldsymbol{\beta}$ in the limit of an infinite trading horizon.

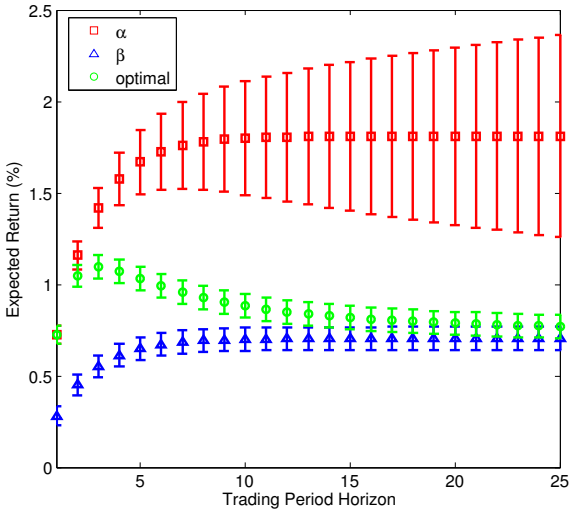


Fig. 3. Expected return as a function of trading horizon, with portfolio standard deviation bars.

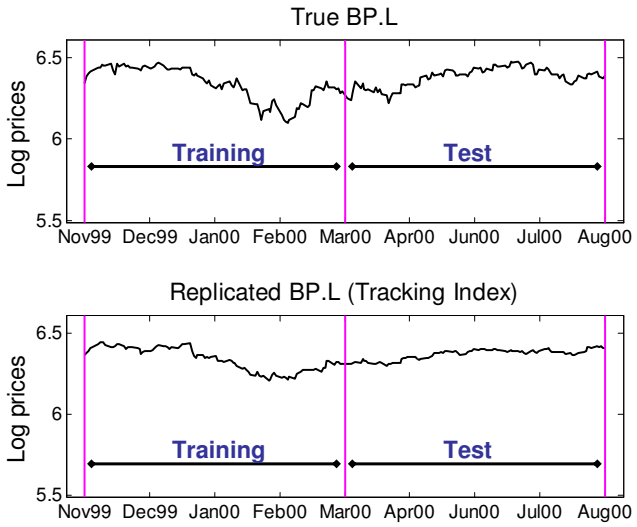


Fig. 4. Daily closing log prices for BP.L and the corresponding synthetic asset.

VI. Experimental Results

In this section we compare the performance of a portfolio constructed using the mean-variance optimal weights given in Eq. (5), with a portfolio whose weights are chosen in the direction of the cointegrating vector, $\boldsymbol{\beta}$, as is commonly done. The dataset from [6], which was chosen as the basis for the experiment, consists of the British Oil (symbol BP.L) stock from the STOXX 50 index from September 14, 1998 to July 3, 2002, and a replicating portfolio, or tracking index, constructed from the remaining 49 assets, chosen to be cointegrated with BP.L. This is a standard construction, done in order to generate a system in which the log-prices are actually cointegrated [5]. In order for the results given in this paper to be applicable, the data under consideration must exhibit the cointegration property with no structural breaks or regime shifts. The reliability of such models is not addressed here, but can be found in [11].

Given the BP.L and tracking index datasets, two consecutive 100-day segments were identified in which the parameters of the VAR model remained relatively constant, denoted as $\mathbf{x}_{\text{train}}$ and $\mathbf{x}_{\text{test}}$. The closing log prices of the real and synthetic asset from November 8, 1999 to March 24, 2000 were used to train the cointegrating VAR model, while the log prices from March 27, 2000 to August 11, 2000 were used to test the trading strategy, as shown in Figure 4. A VAR(1) cointegration model with a constant drift term, was fit to the training data using the ML estimators given in [2], and the corresponding residuals were found to have a correlation coefficient of 0.3625. Using higher order VAR models, a significant improvement in the correlation coefficient was not achieved. Figure 6 contains a scatter plot of the real BP.L and corresponding replicating portfolio, for both the training and test data, along with the ML estimates for the $\boldsymbol{\beta}$, $\boldsymbol{\beta}_\perp$, and $\boldsymbol{\alpha}$ subspaces.

A plot of the signal $z = \boldsymbol{\beta}^T \mathbf{x}_{\text{test}}$ is shown in Figure 5. The signal z is often used as a trade entry indicator due to the fact that it is a measure of how far $\mathbf{x}_{\text{test}}$ is from the $\boldsymbol{\beta}_\perp$ space. This signal measures the current state of disequilibrium, and enables the trader to quickly identify mispricing opportunities. A trade entry threshold of $\gamma = 1.5\sigma_z$ was chosen, and the set of potential entry points, $\mathcal{I}$, was constructed according to $\mathcal{I} = \{i \in \{1, 2, ..., T\} : |z_i| > \gamma\}$, where T is the number of days in the test set. For each day indexed by $\mathcal{I}$, the actual returns were computed as a function of trading horizon, for both the optimal and $\boldsymbol{\beta}$ portfolios. Figure 7 shows the excess return generated by the optimal portfolio over the $\boldsymbol{\beta}$ portfolio, averaged over all of the days in $\mathcal{I}$. In the top plot, $\sigma_0 = 0.05$ for all trading horizons, while in the bottom plot the allowable standard deviation grows linearly with the length of the investment period, i.e. $\sigma_0 = 0.001N$. The allowable standard deviation controls the degree to which the portfolio is leveraged, so

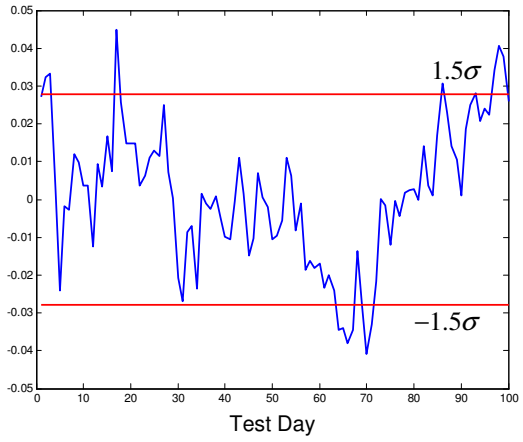
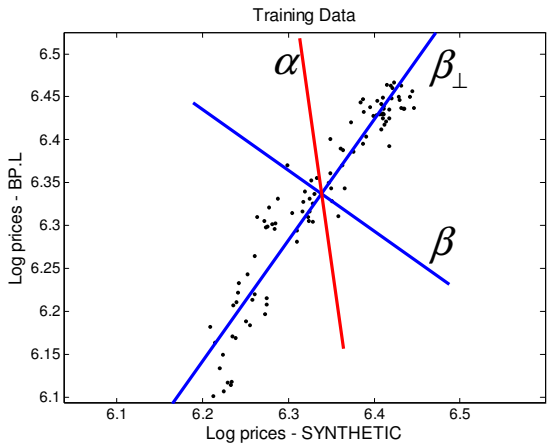


Fig. 5. Trading signal given by stationary signal $\boldsymbol{\beta}^T \mathbf{x}$, using test data. When the threshold is $\gamma = 1.5\sigma_0$, 14 potential entry points were identified.



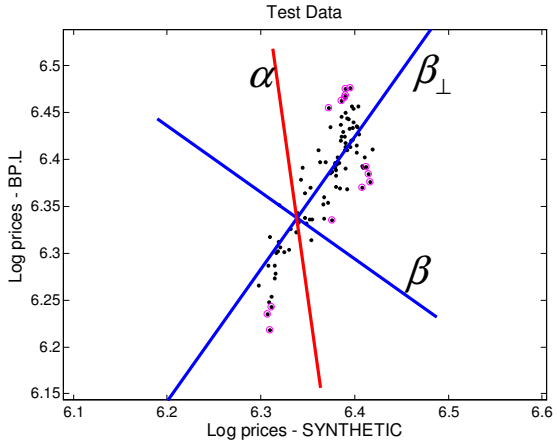
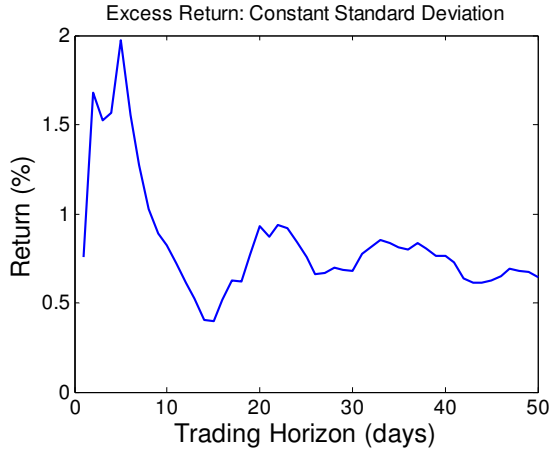


Fig. 6. Scatter plot of the log prices for BP.L and replicating index, along with $\boldsymbol{\beta}$, $\boldsymbol{\beta}_\perp$, and $\boldsymbol{\alpha}$ subspaces. In the test data plot, points surrounded by circles indicate potential entry days.



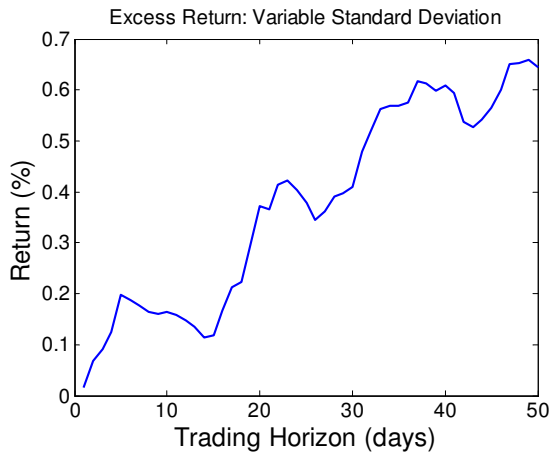


Fig. 7. (top) Excess returns of the optimal portfolio over the $\boldsymbol{\beta}$ portfolio as a function of trading horizon, for fixed variance, averaged over 14 entry points. (bottom) Excess returns, where allowable standard deviation grows linearly with trading horizon length.

that for the variable risk case, the degree of leveraging increases linearly with N. For example, when $N = 1$, the weights indicate that the trader should short sell 1.9 percent of the tracking index, and go long 2.8 percent in BP.L, while for $N = 50$, the trader is instructed to short sell 148 percent of the tracking index and go long 108 percent with BP.L. With constant risk, the degree of leveraging remains relatively uniform for all trading horizons. The largest improvement in return for the constant risk case is realized for short horizons, and the amount of excess return decreases as the optimal portfolio approaches the $\boldsymbol{\beta}$ portfolio in the limit as $N \to \infty$. With variable risk, the degree of improvement initially rises as a function of N, as the amount of allowable leveraging increases, however, the excess return gradually disappears as the mean-variance optimal portfolio converges to the $\boldsymbol{\beta}$ portfolio.

REFERENCES

[1] R. F. Engle and C. W. J. Granger, "Co-integration and error correction: Representation, estimation, and testing," *Econometrica*, vol. 55, no. 2, pp. 251–276, 1987.

[2] S. Johansen, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, New York, 1995.

[3] M. Wagner, "A comparison of johansen's, bierens', and the subspace algorithm method for cointegration analysis," *Oxford Bulletin of Economics and Statistics*, vol. 66, no. 3, pp. 399–424, 2004.

[4] Y. Kawasaki, S. Tachiki, H. Udaka, and T. Hirano, "A characterization of long-short trading strategies based on cointegration," *Proceedings of the 2003 International Conference on Computational Intelligence for Financial Engineering (CIFEr2003)*, pp. 411–416, 2003.

[5] C. Alexander and A. Dimitriu, "Index and statistical arbitrage," *Journal of Portfolio Management*, vol. 31, no. 2, pp. 50–63, 2005.

[6] A. N. Burgess, *Applied Quantitative Methods for Trading and Investment*, chapter Using Cointegration to Hedge and Trade International Equities, Wiley, 2003.

[7] G. W. P. Thompson, "Optimal trading of an asset driven by a hidden markov process in the presence of fixed transaction costs," Centre for Financial Research, Judge Institute of Management, University of Cambridge, 2002.

[8] A. Lucas, "Strategic and tactical asset allocation and the effect of long-run equilibrium relations," Research Memorandum 1997-42, Vrije Universiteit, Amsterdam, August 1997.

[9] C. T. Chen, *Introduction to Linear System Theory*, Holt, Rinhart and Winston, 2002.

[10] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[11] Peter Reinhard Hansen, *Structural Changes in Cointegrated Processes*, Ph.D. thesis, University of California at San Diego, 2000.