

A Dynamic Programming Approach to Two-Stage Mean-Variance Portfolio Selection 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—In this paper we study the problem of optimal portfolio construction when the trading horizon consists of two consecutive decision intervals and rebalancing is permitted. It is assumed that the log-prices of the underlying assets are non-stationary, and specifically follow a discrete-time cointegrated vector autoregressive model. We extend the classical Markowitz mean-variance optimization approach to a multi-period setting, in which the new objective is to maximize the total expected return, subject to a constraint on the total allowable risk. In contrast to traditional approaches, we adopt a definition for risk which takes into account the non-zero correlations between the inter-stage returns. This portfolio optimization problem amounts to not only determining the relative proportions of the assets to hold during each stage, but also requires one to determine the degree of portfolio leverage to assume. Due to a fixed constraint on the standard deviation of the total return, the leverage decision is equivalent to deciding how to optimally partition the allowed variance, and thus variance can be viewed as a shared resource between the stages. We derive the optimal portfolio weights and variance scheduling scheme for a trading strategy based on a dynamic programming approach, which is utilized in order to make the problem computationally tractable. The performance of this method is compared to other trading strategies using both Monte Carlo simulations and real data, and promising results are obtained.

I. INTRODUCTION

It is often stated that many groups of real-world macro-economic variables are cointegrated, meaning they are well modeled by a vector autoregressive process containing at least one common stochastic trend [1]. In these systems, the time series corresponding to the prices of individual assets are nonstationary, while the series of first differences are stationary. In addition, it is possible to construct a linear combination of the signals, i.e. a portfolio, that is stationary, thereby removing the common source of nonstationarity. Given the popularity of this model both in the literature and among practitioners, we address the question of optimal portfolio construction given a universe of cointegrated assets.

The problem of portfolio construction in cointegrated vector autoregressive systems has been previously studied. Early work focused on the use of statistical arbitrage techniques, such as mean-reverting and momentum strategies, for trading a stationary linear combination of cointegrated assets [2]–[4]. More recently, it has been shown

in [5] that these techniques are not optimal in the classical Markowitz mean-variance sense, and that it is possible to achieve a higher average return for the same level of risk by constructing a portfolio that has a component not only in the direction of bounded variance, but also in the direction of expected change.

The optimal asset allocation rule in [5] is derived for the case where there is a single decision interval corresponding to a finite trading horizon with no ability to rebalance; here we extend this analysis to consider the case where rebalancing of the asset holdings is permitted. Attention is restricted to a two-stage scenario, and the Markowitz framework is extended to this setting. Ideally, we seek the portfolio for each stage that maximizes the expected total portfolio return, subject to a fixed constraint on the portfolio risk. We define risk as the variance of the sum of the per-stage returns, rather than the sum of the per-stage variances, so that we may account for the non-zero inter-stage correlations of the returns induced by our cointegration model. However, we show that it is not possible to compute such portfolios exactly, and therefore we consider an approximation based on a dynamic programming (DP) approach.

The organization of this paper is as follows. In Section II, we present the cointegrated VAR model and two-period mean-variance optimization framework. The derivation of the optimal asset allocation rule for each stage using the dynamic programming approach is given in Section III. Simulation results using synthetic data that contrast our solution to existing methods are analyzed in Section IV, followed by a discussion of a trading simulation based on real, historical data in Section V.

II. PROBLEM FORMULATION

Let \mathbf{x}_k be a 2-dimensional random vector representing the log-prices of a set of two assets, that follow a first-order vector autoregressive, VAR(1), process:

$$\mathbf{x}_{k-1} = \mathbf{\Pi}_1 \mathbf{x}_k + \mathbf{\Phi} \mathbf{d}_k + \boldsymbol{\epsilon}_k. \quad (1)$$

Here the 2×2 $\mathbf{\Pi}_1$ matrix encodes the temporal dependence among the component processes of \mathbf{x}_k ; \mathbf{d}_k is a vector of

deterministic inputs, often containing a constant or linear function; Φ is the matrix relating the elements of \mathbf{d} to \mathbf{x} ; and ϵ_k is a 2-dimensional Gaussian random vector with zero mean and variance Ψ that drives the overall process. Note that the states are numbered in decreasing order, so that \mathbf{x}_k denotes the log-prices of the assets at the k^{th} stage from the end, as depicted in Figure 1 for the two-stage case. Throughout this paper, we restrict attention to the first-order VAR case, but extensions to higher-order VAR systems follow naturally by augmenting the state space.

The model given in Eq. 1 is said to exhibit the cointegration property when the matrix defined as $\Pi = \Pi_1 - \mathbf{I}$ is not of full rank. This occurs when the characteristic equation contains a root at unity, possibly endowing each of the underlying time series of \mathbf{x}_k with a random walk component. The matrix Π_1 has one eigenvalue $\lambda_1 = 1$ and the other with the property that $|\lambda_2| < 1$. Since Π is of rank $r < 2$ and $\Pi \neq \mathbf{0}$, it must be true that $r = 1$, and therefore Π can be expressed as the outer product of two 2×1 vectors, as:

$$\Pi = \alpha\beta^T. \quad (2)$$

The data generated from this random process has finite variance along the direction given by β , and diverging variance in the orthogonal direction, denoted as β_{\perp} . The one-dimensional column space of β is commonly referred to as the cointegrating space, while the column space of α is referred to as the space of disequilibrium adjustment forces. It can be shown that for any \mathbf{b} in the span of $\{\beta\}$, $\mathbf{b}^T \mathbf{x}$ is a wide-sense stationary random process [6].

We extend the classical Markowitz mean-variance portfolio optimization approach [7] to a two-period setting, in which the objective is to maximize the total expected return of the portfolio summed across both periods, subject to a single constraint on the variance of the total return at the end, rather than a set of constraints on the per-stage returns. Formally, the optimization problem, P_0 , is given by:

$$\left. \begin{aligned} \mathbf{w}_1^*, \mathbf{w}_2^* &= \arg \max_{\mathbf{w}_1, \mathbf{w}_2} E[\mathbf{w}_1^T \mathbf{r}_1 + \mathbf{w}_2^T \mathbf{r}_2] \\ \text{s.t. } \text{var}[\mathbf{w}_1^T \mathbf{r}_1 + \mathbf{w}_2^T \mathbf{r}_2] &= \sigma_0^2, \end{aligned} \right\} P_0$$

where the per-period vector of individual asset returns, \mathbf{r}_k , is defined as the change in the log prices, as:

$$\mathbf{r}_k = \Delta \mathbf{x}_k = \mathbf{x}_{k-1} - \mathbf{x}_k.$$

The expectation and variance operators are taken with respect to the information available at the starting time, denoted as t_2 . The inner product represented by $\mathbf{w}_k^T \mathbf{r}_k$ denotes the return of the portfolio for stage k . The stages, like the states, are numbered in reverse order, so that \mathbf{w}_k denotes the relative asset holdings in the k^{th} stage from the end. The portfolio weight vector represents the relative percentage of wealth to allocate to each asset, where a positive weight indicates a long position and negative weight denotes a short position. We allow the portfolio at any stage to be leveraged, i.e. the market value of the

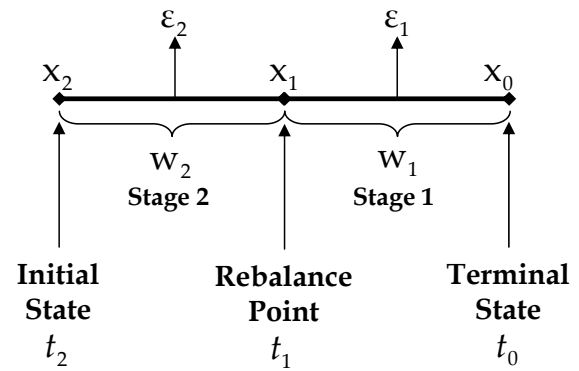


Fig. 1. State sequence for two-stage portfolio optimization problem, with rebalancing. The states (log-prices), times, and stages are numbered in decreasing order, so that the subscript indicates the distance from the terminal point.

portfolio may exceed the available wealth, and therefore a budget constraint of the form $\mathbf{1}^T \mathbf{w}_k = 1$ is not required. The degree of leverage is limited by the allowable risk parameter, σ_0 .

By introducing a Lagrange multiplier, λ , problem P_0 can be rewritten as:

$$\left. \begin{aligned} \mathbf{w}_1^*, \mathbf{w}_2^*, \lambda^* &= \arg \max E[\mathbf{w}_1^T \mathbf{r}_1] + E[\mathbf{w}_2^T \mathbf{r}_2] \\ -\lambda \left\{ \begin{aligned} &\text{var}[\mathbf{w}_1^T \mathbf{r}_1] + \text{var}[\mathbf{w}_2^T \mathbf{r}_2] \\ &+ 2\text{cov}[\mathbf{w}_1^T \mathbf{r}_1, \mathbf{w}_2^T \mathbf{r}_2] - \sigma_0^2 \end{aligned} \right\} \end{aligned} \right\} P'_0$$

At first glance, it appears that an exact solution to P'_0 should be easy to compute. However, the portfolio over the last stage, \mathbf{w}_1 , is itself a random variable, as it depends on the observed value of the state at time t_1 , i.e. $\mathbf{w}_1 = f(\mathbf{x}_1)$. As the nature of this dependence is unknown, it is not possible to immediately compute the terms in P'_0 that depend on \mathbf{w}_1 , whether in closed form or by numerical methods. Furthermore, the problem does not map directly into a dynamic programming context [8], as the mean-variance cost function given in P'_0 is not additive over time due to the non-zero correlation of the per-stage portfolio returns. Additionally, the problem cannot be expressed as the expected utility of the total return due to the presence of the variance operator, which introduces a squared expectation term into the objective function. To address these limitations, we consider a relaxation of problem P'_0 based on the concept of backwards induction from the DP algorithm. First, the optimal portfolio for the last stage is determined to within a scale factor. Once this direction is established, it is possible to solve for both the direction of the second stage from the end and the optimal variance scheduling scheme, resulting in a suboptimal, but computable solution.

III. PORTFOLIO CONSTRUCTION

Here we solve the two-stage portfolio selection problem by applying the dynamic programming backward recursion. We

first consider the tail subproblem consisting of only the last stage, denoted as stage 1 in Figure 1. Looking forward from this time, there is a single decision interval with a holding period of one time step, and therefore we can apply the solution presented in [5] for $N = 1$, yielding:

$$\mathbf{w}_1^* = a_1 \Psi^{-1} \Pi \mathbf{x}_1 = a_1 \mathbf{W}_1 \mathbf{x}_1, \quad (3)$$

where a_1 is a scale factor or degree of leverage to be determined via enforcement of the total variance constraint. Note that the portfolio direction is a linear function of the state, \mathbf{x}_1 , and thus by applying the backwards recursion we have determined a particular form for the function $\mathbf{w}_1 = f(\mathbf{x}_1)$.

We now seek the optimal portfolio for the second to last stage, given our expression for the portfolio for the last stage. For notational simplicity, let:

$$\mathbf{z} = \begin{pmatrix} \mathbf{w}_2^T (\mathbf{x}_1 - \mathbf{x}_2) \\ \mathbf{x}_1^T \mathbf{W}_1^T (\mathbf{x}_0 - \mathbf{x}_1) \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 1 \\ a_1 \end{pmatrix}.$$

The portfolio for the second stage is computed as:

$$\mathbf{w}_2^* = \arg \max_{\mathbf{w}_2} \mathbf{a}^T \boldsymbol{\mu}_z - \lambda \{ \mathbf{a}^T \boldsymbol{\Sigma}_z \mathbf{a} - \sigma_0^2 \}, \quad P_1$$

where $\boldsymbol{\mu}_z$ and $\boldsymbol{\Sigma}_z$ are the mean vector and covariance matrix of \mathbf{z} , respectively, exact expressions for which are derived in Appendix A. The solution to problem P_1 is given by:

$$\mathbf{w}_2^* = \left(\frac{1}{2\lambda} \Psi^{-1} \Pi - a_1 (\Pi^T \mathbf{W}_1 + \mathbf{W}_1^T \Pi) \Pi_1 \right) \mathbf{x}_2. \quad (4)$$

This expression for \mathbf{w}_2 has many interesting properties. First, we observe that Eq. 4 is also a linear function of the current state, and therefore can be rewritten as $\mathbf{w}_2 = \mathbf{W}_2 \mathbf{x}_2$. Next, we note that the first term is proportional to $\Psi^{-1} \Pi \mathbf{x}_2$, which has identical structure to Eq. 3. This component corresponds to a scaled version of the optimal solution for a single stage problem beginning at time t_2 , and thus can be thought of as the ‘‘myopic’’ component. In this light, the second term can be viewed as a correction factor that modifies the myopic solution to account for the uncertainty of the new log-price information, \mathbf{x}_1 , which becomes available at the rebalance time, t_1 . This modification depends both on the direction and scaling of \mathbf{w}_1 , as is evidenced by the explicit presence of both a_1 and \mathbf{W}_1 factors in Eq. 4. As shown in Appendix A, this correction factor results from the non-zero covariance between the components of the random vector \mathbf{z} . In Section IV, we show that this direction modification has the effect of increasing the negative correlation between the returns for stages 1 and 2, enabling an increase in the amount of leverage realized for each period, while maintaining a constant level of total risk.

All that remains is to determine the precise variance scheduling scheme, or per-stage leverage amounts that must be exercised in order to meet the total variance constraint. We seek values for the scale factors a_1 and λ that maximize

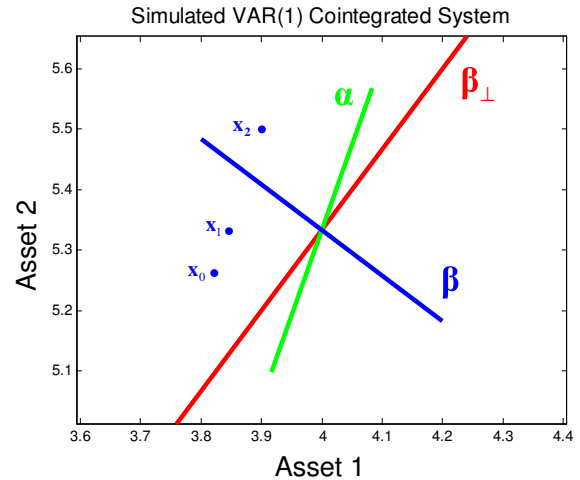


Fig. 2. Geometric view of cointegrated vector autoregressive system in \mathbf{x}_k space. The three points represent a single path of the random process defined by Eq. 5, beginning from \mathbf{x}_2 . The data generated according to this model has infinite variance along the β_{\perp} direction, and finite variance in the β direction. The α vector indicates the direction disequilibrium readjustment forces.

the objective function given in problem P_1 . As derived in Appendix A, we find that:

$$a_1^* = \left(\frac{1}{2\lambda} \right) \frac{E[z_1] - \mathbf{x}_2^T \Pi^T \Psi^{-1} \mathbf{W}_{1,2} \mathbf{x}_2}{\text{var}[z_1] - \mathbf{x}_2^T \mathbf{W}_{1,2}^T \Psi^{-1} \mathbf{W}_{1,2} \mathbf{x}_2}$$

$$\frac{1}{2\lambda} = \frac{\sigma_0}{\sqrt{\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2 + A_1^2 (\text{var}[z_1] - \mathbf{x}_2^T \mathbf{W}_{1,2}^T \Psi^{-1} \mathbf{W}_{1,2} \mathbf{x}_2)}}$$

where \mathbf{A} and $\mathbf{W}_{1,2}$ are defined in Equations 6 and 7. While we have chosen to focus here on the two-stage case for simplicity and clarity, extending to the N stage case follows naturally by augmenting the \mathbf{z} and \mathbf{a} vectors, and continuing to apply the DP backwards recursion.

IV. SIMULATION RESULTS

In order to better understand the portfolio directions and variance scheduling scheme derived in Section III, we consider a representative example using data generated from a synthetic model. We compare the portfolios computed using the DP approach to a set of three existing techniques, given by:

- *The ‘beta’ portfolio*: Here the assets are allocated in the direction given by the β vector from the cointegrated VAR model, defined according to Eq. 2, irrespective of the observed state variables. Rebalancing is prohibited, and the portfolio is scaled in order to meet the variance constraint. This scheme is commonly used by practitioners, and is the basis for a wide variety of statistical arbitrage techniques. For additional details, see [2].
- *The ‘Markowitz, without rebalancing’ portfolio*: The asset allocation rule is formed by considering a single decision interval of length $N = 2$, and applying the result from [5] for the optimal mean-variance portfolio in a cointegrated VAR system.
- *A ‘semi-myopic’ portfolio*: The result from [5] is independently applied over two consecutive intervals, in

order to determine the portfolio *directions* for each stage. Next, these vectors are appropriately scaled so that the total variance constraint is maintained. The name highlights the fact that the directions are chosen myopically, while the scale factors are not. Additional details are provided in Appendix B.

We first contrast the behavior of each trading strategy by examining the second-order statistics of the per-stage and total returns, computed via Monte Carlo simulations. This is followed by a comparison of the four asset allocation schemes using a single, representative sample path.

Consider the following synthetic VAR(1) model, with no deterministic inputs:

$$\mathbf{x}_{k-1} = \begin{pmatrix} 1.18 & -0.14 \\ 0.51 & 0.62 \end{pmatrix} \mathbf{x}_k + \boldsymbol{\epsilon}_k, \quad (5)$$

where $\boldsymbol{\epsilon}_k \sim N(\mathbf{0}, \boldsymbol{\Psi})$ and $\boldsymbol{\Psi} = 0.001\mathbf{I}$. In this system, $\boldsymbol{\alpha} = (-0.28 \ -0.77)^T$ and $\boldsymbol{\beta} = (-0.66 \ 0.5)^T$, as depicted in Figure 2. The initial log-price pair for all of the simulations was chosen to be $\mathbf{x}_2 = (3.9 \ 5.5)^T$, and we are interested in determining the optimal portfolio weights in all four trading scenarios for the case where the total level of the allowed risk is given by $\sigma_0 = 0.05$, or 5%.

The system in Eq. 5 is simulated $M=10^4$ times, and the resulting per-stage and total return statistics are given in Table I. The table also displays the correlation coefficient of the inter-stage returns, and it is here that we begin to gain some intuition for the DP solution. As compared to the other approaches, the weights derived via the DP approach achieve a higher negative correlation between the per-period returns, which enables the per-stage variances to be greater in magnitude in contrast to alternative algorithms. In fact, the per-stage variances are each greater than σ_0^2 , while the negative correlation among per-stage returns enables the total variance constraint to still be met, resulting in a higher expected return.

Figure 2 illustrates one sample path generated from Eq. 5. The resulting portfolio directions are illustrated in Figure 3, while the exact leverage amounts are presented in Table II. The table also displays the total return achieved by each strategy for this particular sample path. We find that the degree of leverage utilized in the DP approach is greater than all other strategies, which is the main source of the increased realized return.

V. EXPERIMENTAL RESULTS

Here we compare the performance of the dynamic programming trading strategy of Section III with the three strategies described in Section IV, using historical price data. The selected dataset from [4] consists of the British Oil (symbol BPL) stock from the STOXX 50 index, and a replicating portfolio, or tracking index, constructed from the remaining 49 assets, so that the two series exhibit the cointegration property, with no structural breaks or regime

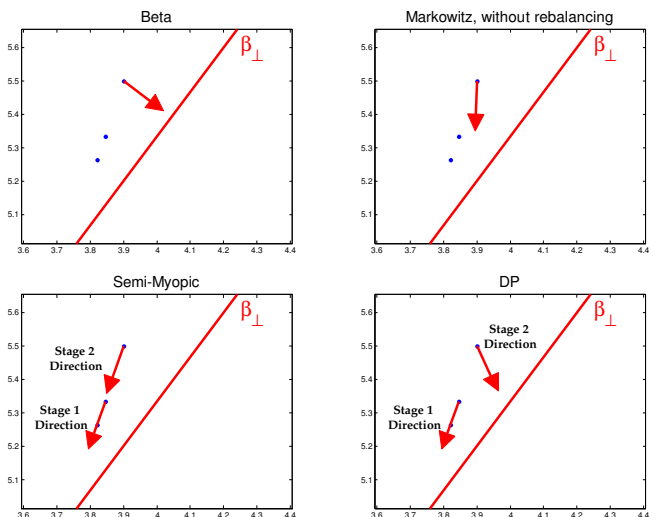


Fig. 3. Comparison of portfolio directions in the \mathbf{x}_k space for all four trading strategies. Since rebalancing is prohibited in the beta and Markowitz schemes, only a single arrow is shown.

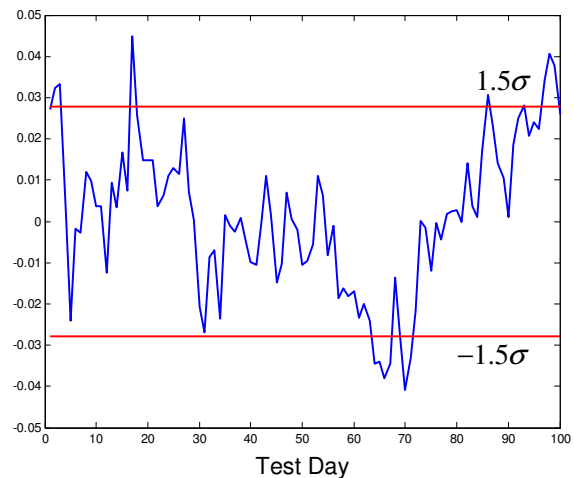


Fig. 4. Stationary trading indicator signal, $z = \boldsymbol{\beta}^T \mathbf{x}$, used to determine when to enter into test portfolios. Portfolios are bought when $|z| > 1.5\sigma_z$ and are sold two periods later.

shifts [3].

Given the BPL and tracking index datasets, two consecutive 100-day data segments, denoted as $\mathbf{x}_{\text{train}}$ and \mathbf{x}_{test} , were identified in which the parameters of the VAR model remained relatively constant. The closing log prices from November 8, 1999 to March 24, 2000 were used to train the cointegrated VAR model, while the log prices from March 27, 2000 to August 11, 2000 were used to test the trading strategy. A VAR(1) cointegration model with a constant drift term was fit to the training data using the ML estimators described in [6]. A significant decrease in correlation coefficient of the residuals was not achieved by considering higher-order VAR models.

Trading Strategy	Stage 2		Stage 1		Correlation	Total	
	Mean	Var	Mean	Var		Mean	Var
Beta	0.045	0.0015	0.036	0.0016	-0.19	0.081	0.0025
Markowitz, without rebalancing	0.14	0.0015	0.11	0.0021	-0.31	0.25	0.0025
Semi-Myopic	0.18	0.002	0.06	0.0009	-0.24	0.24	0.0025
DP	0.18	0.0045	0.12	0.0039	-0.70	0.30	0.0025

TABLE I

SECOND ORDER STATISTICS OF RETURNS FROM MONTE CARLO SIMULATIONS.

Trading Strategy	Stage 2 Leverage		Stage 1 Leverage		Total Return
	Asset 1	Asset 2	Asset 1	Asset 2	
Beta	0.99	-0.74	0.99	-0.74	0.10
Markowitz, without rebalancing	-0.04	-1.22	-0.04	-1.22	0.29
Semi-Myopic	-0.50	-1.41	-0.16	-0.46	0.30
DP	0.91	-1.90	-0.33	-0.93	0.34

TABLE II

ACTUAL PORTFOLIO LEVERAGE (% OF INITIAL WEALTH) VALUES FOR SINGLE SAMPLE PATH, WITH $\sigma_0 = 0.05$.

Trading Strategy	Stage 2		Stage 1		Total	
	Mean	Var	Mean	Var	Mean	Var
Beta	0.0207	0.0009	0.0190	0.0006	0.0397	0.0018
Markowitz, without rebalancing	0.0252	0.0014	0.0208	0.0011	0.0460	0.0030
Semi-Myopic	0.0278	0.0020	0.0346	0.0027	0.0624	0.0056
DP	0.0275	0.0016	0.0357	0.0028	0.0631	0.0054

TABLE III

SECOND ORDER STATISTICS OF RETURNS FROM REAL DATA EXAMPLE.

The trading strategy implemented using this dataset works as follows. For each data point in the test set, we compute a test statistic, $z = \beta^T \mathbf{x}$, as shown in Figure 4. When $|z| > 1.5\sigma_z$, a decision is made to “enter the market”, here resulting in 8 entry points. The portfolio weights for the next two days (stages) are computed according to each strategy using $\sigma_0 = 0.05$. The per-stage and total return statistics are displayed in Table III. We note that the total variances reported in Table III are not equal to σ_0^2 , which is due not only to the small sample size but also the fact that we are averaging over initial values of \mathbf{x}_2 . We observe that all of the approaches achieved an average return in the second stage from the end between two and three percent. However, in the last stage from the end, the DP and semi-myopic strategies beat the two non-rebalancing strategies by over one percent, due to the fact that they take advantage of the new log-price information that becomes available at the rebalance point. As a result of this truly dynamic trading methodology, these strategies achieve a higher total return for each initial condition, while maintaining a constant level of total risk. As we saw in the Monte Carlo simulations of Section IV, it is the DP strategy that is able to achieve the highest expected return, due to the increase in the negative correlation of the inter-stage returns.

APPENDIX A

In this Appendix we derive expressions for μ_z , Σ_z , \mathbf{w}_2 , a_1 , and λ . We begin with μ_z , and recall that all expectations are computed with respect to the information available at the

beginning of the second to last stage, time t_2 . Let

$$\mathbf{z} = \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} = \begin{pmatrix} \mathbf{w}_2^T (\mathbf{x}_1 - \mathbf{x}_2) \\ \mathbf{x}_1^T \mathbf{W}_1^T (\mathbf{x}_0 - \mathbf{x}_1) \end{pmatrix},$$

and we have:

$$\begin{aligned} E[z_2] &= E[\mathbf{w}_2^T (\mathbf{x}_1 - \mathbf{x}_2)] \\ &= \mathbf{w}_2^T E[\mathbf{\Pi} \mathbf{x}_2 + \boldsymbol{\epsilon}_2] \\ &= \mathbf{w}_2^T \mathbf{\Pi} \boldsymbol{\mu}_2, \\ E[z_1] &= E[\mathbf{x}_1^T \mathbf{W}_1^T (\mathbf{x}_0 - \mathbf{x}_1)] \\ &= E[\mathbf{x}_1^T \mathbf{W}_1^T \mathbf{\Pi} \mathbf{x}_1 + \mathbf{x}_1^T \mathbf{W}_1^T \boldsymbol{\epsilon}_1] \\ &= E[\mathbf{x}_1^T \mathbf{W}_1^T \mathbf{\Pi} \mathbf{x}_1] \\ &= \mathbf{x}_2^T \mathbf{\Pi}_1^T \mathbf{W}_1^T \mathbf{\Pi} \mathbf{\Pi}_1 \mathbf{x}_2 + \text{trace}[\mathbf{W}_1^T \mathbf{\Pi} \boldsymbol{\Psi}], \end{aligned}$$

where $\mathbf{W}_1 = \boldsymbol{\Psi}^{-1} \mathbf{\Pi}$. We now compute each of the terms in Σ_z . The variance of z_2 is easily computed as:

$$\begin{aligned} \text{var}[z_2] &= \text{var}[\mathbf{w}_2^T (\mathbf{x}_1 - \mathbf{x}_2)] \\ &= \mathbf{w}_2^T \boldsymbol{\Psi} \mathbf{w}_2. \end{aligned}$$

In order to compute the variance of z_1 , we invoke the law of total variance, as:

$$\begin{aligned} \text{var}[z_1] &= \text{var}[E[z_1 | \mathbf{x}_1]] + E[\text{var}[z_1 | \mathbf{x}_1]] \\ &= \text{var}[\mathbf{x}_1^T \mathbf{W}_1^T \mathbf{\Pi} \mathbf{x}_1] + E[\mathbf{x}_1^T \mathbf{W}_1^T \boldsymbol{\Psi} \mathbf{W}_1 \mathbf{x}_1]. \end{aligned}$$

We now define the symmetric matrix \mathbf{A} as:

$$\mathbf{A} \triangleq \mathbf{W}_1^T \mathbf{\Pi} = \mathbf{\Pi}^T \boldsymbol{\Psi}^{-1} \mathbf{\Pi} = \mathbf{W}_1^T \boldsymbol{\Psi} \mathbf{W}_1, \quad (6)$$

Accordingly, we can express the variance of z_1 as:

$$\begin{aligned}\text{var}[z_1] &= \text{var}[\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1] + E[\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1] \\ &= 4\mathbf{x}_2^T \mathbf{\Pi}_1^T \mathbf{A} \mathbf{\Psi} \mathbf{A} \mathbf{\Pi}_1 \mathbf{x}_2 + 2\text{trace}[\mathbf{A} \mathbf{\Psi} \mathbf{A} \mathbf{\Psi}] \\ &\quad + \mathbf{x}_2^T \mathbf{\Pi}_1^T \mathbf{A} \mathbf{\Pi}_1 \mathbf{x}_2 + 2\text{trace}[\mathbf{A} \mathbf{\Psi}]\end{aligned}$$

Lastly, the covariance is computed as:

$$\begin{aligned}\text{cov}[z_1, z_2] &= E[z_2 z_1] - E[z_2] E[z_1] \\ &= E[\mathbf{w}_2^T (\mathbf{\Pi} \mathbf{x}_2 + \boldsymbol{\epsilon}_2) z_1] - E[z_2] E[z_1] \\ &= \mathbf{w}_2^T E[\boldsymbol{\epsilon}_2 z_1] \\ &= \mathbf{w}_2^T E[\boldsymbol{\epsilon}_2 \mathbf{x}_1^T \mathbf{W}_1^T (\mathbf{\Pi} \mathbf{x}_1 + \boldsymbol{\epsilon}_1)] \\ &= \mathbf{w}_2^T E[\boldsymbol{\epsilon}_2 \mathbf{x}_1^T \mathbf{W}_1^T \mathbf{\Pi} \mathbf{x}_1] \\ &= \mathbf{w}_2^T E[\boldsymbol{\epsilon}_2 (\mathbf{\Pi}_1 \mathbf{x}_2 + \boldsymbol{\epsilon}_2)^T \mathbf{W}_1^T \mathbf{\Pi} (\mathbf{\Pi}_1 \mathbf{x}_2 + \boldsymbol{\epsilon}_2)] \\ &= \mathbf{w}_2^T \mathbf{\Psi} (\mathbf{\Pi}^T \mathbf{W}_1 + \mathbf{W}_1^T \mathbf{\Pi}) \mathbf{\Pi}_1 \mathbf{x}_2 \\ &= \mathbf{w}_2^T \mathbf{W}_{1,2} \mathbf{x}_2.\end{aligned}\quad (7)$$

Now that we have all of the terms in $\boldsymbol{\mu}_z$ and $\boldsymbol{\Sigma}_z$, we can compute \mathbf{w}_2^* by differentiating the objective function in Problem P_1 with respect to w_2 , as:

$$\begin{aligned}0 &= \mathbf{\Pi} \mathbf{x}_2 - \lambda 2 \mathbf{\Psi} \mathbf{w}_2 - 2\lambda a_1 \mathbf{W}_{1,2} \mathbf{x}_2 \\ \mathbf{w}_2^* &= \frac{1}{2\lambda} \mathbf{\Psi}^{-1} (\mathbf{\Pi} \mathbf{x}_2 - 2\lambda a_1 \mathbf{W}_{1,2} \mathbf{x}_2) \\ &= \left(\frac{1}{2\lambda} \mathbf{\Psi}^{-1} \mathbf{\Pi} - a_1 (\mathbf{\Pi}^T \mathbf{W}_1 + \mathbf{W}_1^T \mathbf{\Pi}) \mathbf{\Pi}_1 \right) \mathbf{x}_2.\end{aligned}$$

The scale factor applied to the last stage can be found by differentiating the objective function in Problem P_1 with respect to a_1 , as:

$$\begin{aligned}0 &= E[z_1] - 2\lambda a_1 \text{var}[z_1] - 2\lambda \text{cov}[z_1, z_2] \\ &= E[z_1] - 2\lambda a_1 \text{var}[z_1] - 2\lambda \mathbf{w}_2^T \mathbf{W}_{1,2} \mathbf{x}_2 \\ &= E[z_1] - 2\lambda a_1 \text{var}[z_1] - \mathbf{x}_2^T \mathbf{\Pi}^T \mathbf{\Psi}^{-1} \mathbf{W}_{1,2} \mathbf{x}_2 \\ &\quad + 2\lambda a_1 \mathbf{x}_2^T \mathbf{\Pi}_1^T (\mathbf{\Pi}^T \mathbf{W}_1 + \mathbf{W}_1^T \mathbf{\Pi}) \mathbf{W}_{1,2} \mathbf{x}_2 \\ a_1^* &= \left(\frac{1}{2\lambda} \right) \frac{E[z_1] - \mathbf{x}_2^T \mathbf{\Pi}^T \mathbf{\Psi}^{-1} \mathbf{W}_{1,2} \mathbf{x}_2}{\text{var}[z_1] - \mathbf{x}_2^T \mathbf{W}_{1,2}^T \mathbf{\Psi}^{-1} \mathbf{W}_{1,2} \mathbf{x}_2} \\ &= \frac{1}{2\lambda} A_1\end{aligned}$$

Finally, the value of the quantity $\frac{1}{2\lambda}$ is found to be:

$$\begin{aligned}\sigma_0^2 &= \mathbf{w}_2^T \mathbf{\Psi} \mathbf{w}_2 + a_1^2 \text{var}[z_1] + 2a_1 \mathbf{w}_2^T \mathbf{W}_{1,2} \mathbf{x}_2 \\ &= \left(\frac{1}{2\lambda} \right)^2 \left[\mathbf{x}_2^T (\mathbf{\Pi} - A_1 \mathbf{W}_{1,2})^T \mathbf{\Psi}^{-1} (\mathbf{\Pi} - A_1 \mathbf{W}_{1,2}) \mathbf{x}_2 \right. \\ &\quad \left. + A_1^2 \text{var}[z_1] + 2A_1 \mathbf{x}_2^T (\mathbf{\Pi} - A_1 \mathbf{W}_{1,2})^T \mathbf{\Psi}^{-1} \mathbf{W}_{1,2} \mathbf{x}_2 \right] \\ \frac{1}{2\lambda} &= \frac{\sigma_0}{\sqrt{\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2 + A_1^2 (\text{var}[z_1] - \mathbf{x}_2^T \mathbf{W}_{1,2}^T \mathbf{\Psi}^{-1} \mathbf{W}_{1,2} \mathbf{x}_2)}},\end{aligned}$$

where \mathbf{A} is defined according to Eq. 6.

APPENDIX B

Here we present the problem formulation and solution for the semi-myopic approach. The two stage problem is solved as two consecutive one stage problems, in which the direction of the portfolio for each stage is selected to be equal to the optimal action for a single stage problem, with no consideration given to past or future stages. Once these directions are computed, the degree of leverage is determined so that the total expected return is maximized while ensuring that the variance constraint is met. Applying the approach in [5] independently for each period, we have:

$$\begin{aligned}\mathbf{w}_2^* &= a_2 \mathbf{\Psi}^{-1} \mathbf{\Pi} \mathbf{x}_2, \\ \mathbf{w}_1^* &= a_1 \mathbf{\Psi}^{-1} \mathbf{\Pi} \mathbf{x}_1,\end{aligned}$$

where the a_k 's are scale factors that determine the degree of leverage of the portfolio at stage k . These factors are determined by solving problem P_2 , as:

$$a_1^*, a_2^* = \arg \max_{a_1, a_2} \mathbf{a}'^T \boldsymbol{\mu}_{z'} - \lambda \left\{ \mathbf{a}'^T \boldsymbol{\Sigma}_{z'} \mathbf{a}' = \sigma_0^2 \right\}, \quad \} P_2$$

where:

$$\begin{aligned}\mathbf{z}' &= \begin{pmatrix} \mathbf{x}_2^T \mathbf{\Pi}^T \mathbf{\Psi}^{-1} (\mathbf{x}_1 - \mathbf{x}_2) \\ \mathbf{x}_1^T \mathbf{\Pi}^T \mathbf{\Psi}^{-1} (\mathbf{x}_0 - \mathbf{x}_1) \end{pmatrix}, \\ \mathbf{a}' &= \begin{pmatrix} a_2 \\ a_1 \end{pmatrix},\end{aligned}$$

and $\boldsymbol{\mu}_{z'}$ and $\boldsymbol{\Sigma}_{z'}$ refer to the mean vector and covariance matrix of \mathbf{z}' , respectively. The optimal scale factors are:

$$\begin{aligned}\begin{pmatrix} a_2^* \\ a_1^* \end{pmatrix} &= \frac{1}{2\lambda} \boldsymbol{\Sigma}_{z'}^{-1} \boldsymbol{\mu}_{z'} \\ \frac{1}{2\lambda} &= \frac{\sigma_0}{\sqrt{\boldsymbol{\mu}_{z'}^T \boldsymbol{\Sigma}_{z'}^{-1} \boldsymbol{\mu}_{z'}}}.\end{aligned}$$

REFERENCES

- [1] J. Stock and M. Watson, "Testing for common trends," *Journal of the American Statistical Association*, vol. 83, no. 404, pp. 1097–1107, 1988.
- [2] 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.
- [3] C. Alexander and A. Dimitriu, "Index and statistical arbitrage," *Journal of Portfolio Management*, vol. 31, no. 2, pp. 50–63, 2005.
- [4] A. N. Burgess, *Applied Quantitative Methods for Trading and Investment*, chapter Using Cointegration to Hedge and Trade International Equities, Wiley, 2003.
- [5] M. Rudoy and C. Rohrs, "Optimal portfolio construction in cointegrated vector autoregressive systems," *Proceedings of the 2008 American Control Conference (ACC)*, 2008.
- [6] S. Johansen, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, New York, 1995.
- [7] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [8] D. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, MA, 2000.