DynOpt: Incorporating Dynamics into Mean-Variance Portfolio Optimization

Marco Signoretto
ESAT-SCD/SISTA
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10, B-3001 Leuven (BELGIUM)
Email: marco.signoretto@esat.kuleuven.be

Johan A. K. Suykens
ESAT-SCD/SISTA
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10, B-3001 Leuven (BELGIUM)
Email: johan.suykens@esat.kuleuven.be

***Abstract*—Mean-variance (MV) portfolio theory leads to relatively simple and elegant numerical problems. Nonetheless, the approach has been criticized for treating the market parameters as if they were constant over time. We propose a novel convex optimization problem that extends an existing MV formulation with chance constraint(s) by accounting for the portfolio dynamics. The core idea is to consider a multiperiod scenario where portfolio weights are implicitly regarded as the output of a state-space dynamical system driven by external inputs. The approach leverages a result on realization theory and uses the nuclear norm to penalize complex dynamical behaviors. The proposed ideas are illustrated by two case studies.**

I. Introduction

A. Motivation and Related Work

Portfolio optimization is concerned with allocating fractions of wealth to a set of prescribed assets. Mean-variance (MV) portfolio theory approaches the problem based on a risk-return trade-off; it postulates that risk-adverse investors should only be interested in portfolios with minimal risk. Markowitz phrased asset allocation as a simple and elegant mathematical problem; he showed that, given a lower bound on the return the investor is willing to accept, the optimal portfolio can be obtained by solving a convex quadratic programming problem. On the one hand, the approach can be easily applied and analyzed. On the other hand, the result depends heavily on how the assets' risk and return are estimated and it is therefore sensitive to estimation errors [2], [5]. Different methods have been proposed to address this shortcoming. One approach is to account for uncertainties on the market parameters; in particular, [13], [12] studied alternative problem formulations within the framework of robust convex optimization. Another way to improve the approach due to Markowitz is by taking into account the structure of returns. Many research papers reported that the co-movement of returns of certain financial assets are partially explained by a small number of factors [21], [14], [9]. This knowledge has been leveraged in a number of market-aware asset allocation strategies; [8], for example, developed and tested a methodology based on Kalman filtering that combines information from a MV optimization technique along with a three factor regression model. A factor-based asset return model is also considered in [25], which tackles portfolio optimization by a receding horizon control approach.

B. Main Contribution

The main contribution of the paper is to propose a novel convex optimization problem for portfolio allocation. The approach leverages a result on realization theory of state-space dynamical systems and uses the nuclear norm to penalize complex dynamical behaviors. This leads us to extend the framework for mean-variance portfolio allocation with Value at Risk (VaR) constraint(s) by taking into account the portfolio dynamics. The core idea is to assume that portfolio weights are outputs of an underlying dynamical systems. The behavior of the system is determined by imposing that present portfolio weights should satisfy a risk-return trade-off while ensuring that past portfolio weights maximize the realized (past) returns. In experiments we use as inputs factors that are believed to partially explain the behavior of asset returns.

C. Organization of the Paper

The rest of the paper is organized as follows. In Section II we recall elements of MV portfolio optimization as well as instrumental properties of state-space models. In Section III we present and discuss our novel problem formulation. In Section IV we validate our approach in case-studies. We end the paper with Section V by drawing our concluding remarks.

II. Background

A. Notation

For a positive integer P we write $\mathbb{N}_P$ to denote the set $\{1,\dots,P\}$. For a generic matrix $A \in \mathbb{R}^{D_1 \times D_2}$ we denote by $a_{d_1 d_2}$ the entry indexed by (d_1, d_2); we write a_{d_1} to mean the d_1-th column. For a vector $r \in \mathbb{R}^D$ we denote by $\mathrm{diag}(r)$ the $D \times D$ diagonal matrix satisfying $\mathrm{diag}(r)_{dd} = r_d$, $d \in \mathbb{N}_D$. For arbitrary matrices[1] $A, B \in \mathbb{R}^{D_1 \times D_2}$ the (canonical) inner product is defined as:

$$\langle A, B \rangle = \mathrm{trace}(A^\top B) = \sum_{d_1 \in \mathbb{N}_{D_1}} \sum_{d_2 \in \mathbb{N}_{D_2}} a_{d_1 d_2} b_{d_1 d_2} \ . \quad (1)$$

The corresponding Hilbert-Frobenius norm is $\|A\| = \sqrt{\langle A, A \rangle}$. Finally, we recall that if A has rank R, then the singular value decomposition (SVD) of A is

$$A = U \Sigma V^\top, \quad \Sigma = \mathrm{diag}(\sigma) \quad (2)$$

[1] In particular, vectors.

where $U \in \mathbb{R}^{D_1 \times R}$ and $V \in \mathbb{R}^{D_2 \times R}$ are matrices with orthonormal columns, and the vector of singular values σ satisfy

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_R > 0 \ .$$

B. Mean-Variance Portfolio Optimization

Classical portfolio allocation deals with N financial assets held over a period of time, indexed by t. A portfolio is a vector $w_t \in \mathbb{R}^N$ with w_{nt} being the fraction of wealth invested in asset n throughout period t; p_{nt} is the asset price at t; the one-period net return of the $n-$th asset measured at the beginning of period $t+1$ is

$$z_{n(t+1)} = (p_{n(t+1)} - p_{nt})/p_{nt} \ . \tag{3}$$

Correspondingly, the one-period net return at $t+1$ of the portfolio is

$$r_{t+1} = \sum_{n \in \mathbb{N}_N} w_{nt} z_{n(t+1)} = w_t^\top z_{t+1} \ .$$

Mean-Variance Portfolio theory assumes random changes in prices: z_{t+1} is a random vector with mean m_{t+1} and covariance Σ_{t+1}. Therefore the return corresponding to the portfolio w_t is a random variable (r.v.) with mean $w_t^\top m_{t+1}$ and variance $w_t^\top \Sigma_{t+1} w_t$.

1) Markowitz portfolio optimization: Given a lower bound $r_{\min}$ on the expected return the investor is willing to accept in $T+1$, MV portfolio optimization is concerned with finding a portfolio $\hat{w}_T$ with minimal return variance. The classical approach due to Markowitz prescribes to solve the quadratic programming (QP) problem[2] [18]:

$$\begin{array}{ll} \min\limits_{w_T} & w_T^\top \hat{\Sigma} w_T \\ \text{subject to} & w_T^\top \hat{m} \geq r_{\min} \\ & 1^\top w_T = 1; \ w_{nT} \geq 0, \ n \in \mathbb{N}_N \end{array} \tag{MkwOpt}$$

where $\hat{\Sigma}$ and $\hat{m}$ are the *market parameters*, i.e., estimates for respectively, Σ_{T+1} and m_{T+1}. In (MkwOpt), $1^\top w_T = 1$ is a budget constraint and $w_{Tn} \geq 0, \ n \in \mathbb{N}_N$ impose long-only positions[3].

2) Portfolio Optimization with VaR Constraints: A possible extension to the standard MV framework is to include VaR constraints [13]. A VaR constraint imposes that an upper bound[4] α on the expected return should hold with probability smaller than β:

$$\text{prob}(w_T^\top \hat{m} \leq \alpha) \leq \beta \ . \tag{4}$$

[2]In problem (MkwOpt) and, later, in (VaROpt), we denoted the portfolios as if they were indexed by T, the present time period. This is not common in the literature since these problems are concerned with single-period investments and regard the market parameters and — consequently — the portfolio, as if they were constant over time. We adopted this notation for the sake of clarity to contrast these problems with problem (DynOpt) in Section III, where one considers a multiperiod scenario comprising also the past portfolios.

[3]Throughout this paper we consider long-only portfolios; nevertheless, the proposed approach can be extended to the case where short positions are allowed.

[4]Here α is an undesired — typically negative — level of expected return.

In the following we refer to α and β as the *VaR parameters*. If m_u and σ_u^2 denote respectively the mean and the variance of the r.v. $u = w_T^\top z_{T+1}$, (4) can be restated as

$$\text{prob}\left(\frac{u - m_u}{\sigma_u} \leq \frac{\alpha - m_u}{\sigma_u} \right) \leq \beta \ . \tag{5}$$

Under gaussianity assumptions on the returns, $(u - m_u)/\sigma_u$ is normal. It follows that (5) can be expressed as $\frac{\alpha - m_u}{\sigma_u} \leq \Phi^{-1}(\beta)$ where Φ denotes the cumulative distribution function of a normally distributed r.v.. Letting

$$m_u = w_T^\top \hat{m} \quad \text{and} \quad \sigma_u = \left(w_T^\top \hat{\Sigma} w_T \right)^{1/2} \ ,$$

one obtains:

$$w_T^\top \hat{m} + \Phi^{-1}(\beta) \| \hat{\Sigma}^{1/2} w_T \| \geq \alpha \ . \tag{6}$$

We assume that $\beta \leq 1/2$, in which case $\Phi^{-1}(\beta) \leq 0$ and (6) is a *second-order cone constraint*. In this situation, the problem of maximizing the expected return subject to a VaR constraint (4) can be casted as the *second order cone program* (SOCP) [4]:

$$\begin{array}{ll} \max\limits_{w_T} & w_T^\top \hat{m} \\ \text{subject to} & w_T^\top \hat{m} + \Phi^{-1}(\beta) \| \hat{\Sigma}^{1/2} w_T \| \geq \alpha \\ & w_{nT} \geq 0, \ n \in \mathbb{N}_N \\ & 1^\top w_T = 1 \ . \end{array} \tag{VaROpt}$$

SOCPs constitute a special class of convex programs that can be solved efficiency, for instance by interior point methods; since second-order cone constraints can be written as Linear Matrix Inequalities (LMI), SOCPs can be reformulated as semi definite programs (SDPs). Problem (MkwOpt) is also called *stochastic programming problem* since (4) is a probabilistic constraint (a.k.a. *chance constraint*). Finally note that, although in (MkwOpt) we considered only one probabilistic constraint, it is possible to formulate an SOCP problem with multiple constraints each with different return bound and corresponding probability.

C. Realization of State-Space Systems

The approach that we present in the next Section regards the portfolio $w_t \in \mathbb{R}^N$ as the output of a state-space linear time invariant system driven by an external input $u_t \in \mathbb{R}^M$ [17]:

$$\begin{array}{rcl} x_{t+1} & = & Ax_t + Bu_t \\ w_t & = & Cx_t + Du_t \ . \end{array} \tag{7}$$

In the latter, A, B, C, D are the system matrices with the appropriate dimensions and $x_t \in \mathbb{R}^S$ is the latent state vector, with S being the order of the system. State-space systems are a powerful way to model the behavior of dynamical systems. They find applications in the most diverse domains ranging from engineering to neuroscience and econometrics. In the remainder of this section we recall a realization property that is instrumental to our approach.

1) Hankel Equation: For an arbitrary $D \times T$ matrix $V = [v_1, v_2, \ldots, v_T]$, consider the linear block-Hankel operator

$\mathcal{H}_R : \mathbb{R}^{D\times T} \to \mathbb{R}^{DR\times(T-R+1)}$, defined by:

$$\mathcal{H}_R(V) = \begin{bmatrix} v_1 & v_2 & v_3 & \cdots & v_{T-R+1} \\ v_2 & v_3 & v_4 & \cdots & v_{T-R+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_R & v_{R+1} & v_{R+2} & \cdots & v_T \end{bmatrix}. \quad (8)$$

With reference to (7), assume that inputs and outputs measurements are collected within matrices

$$U = [u_1, \ldots, u_T] \in \mathbb{R}^{M\times T} \text{ and } W = [w_1, \ldots, w_T] \in \mathbb{R}^{N\times T}.$$

It is not difficult to see that the recursive equations (7) lead to the matrix equation:

$$\mathcal{H}_R(W) = \Gamma X + H\mathcal{H}_R(U), \quad (9)$$

where

$$\Gamma = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{R-1} \end{bmatrix} = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \vdots \\ \Gamma_R \end{bmatrix}, \; X = \begin{bmatrix} x_1, & x_2, & \ldots, & x_{T-R} \end{bmatrix} \quad (10)$$

and

$$H = \begin{bmatrix} D & 0 & 0 & \ldots & 0 \\ CB & D & 0 & \ldots & 0 \\ CAB & CB & D & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{R-1}B & CA^{R-2}B & CA^{R-3}B & \ldots & D \end{bmatrix}. \quad (11)$$

2) Realization Property: For a generic matrix V, denote by $V^\perp$ the matrix whose columns form a basis for the null space $\mathcal{N}(V) = \{a \; : \; Va = 0\}$. Let us define now:

$$U_H^\perp = (\mathcal{H}_R(U))^\perp \in \mathbb{R}^{(T-R+1)\times L} \quad (12)$$

where L is the dimension of $\mathcal{N}(\mathcal{H}_R(U))$. By right-multiplying both sides of (9) by $U_H^\perp$ one obtains:

$$\mathcal{H}_R(W)U_H^\perp = \Gamma X U_H^\perp \; . \quad (13)$$

It can be shown, under mild conditions on U and X, that this matrix has rank S equal to the order of the system, i.e., the dimensionality of the state vector, see [28, Section 3.3]. Since $\mathcal{H}_R(W)U_H^\perp$ does not depends either upon the system matrices or the states, it follows that one can estimate the order of the system solely based on input-output data. This fact has found applications in subspace identification methods, see, e.g., [7], [26].

III. Proposed Approach

A. Complexity of Dynamical Systems

Our interest in the aforementioned realization property arises from the interpretation of the system order as a complexity measure for dynamical systems. In line with the Occam's razor, the principle that states a preference for simple theories, our problem formulation requires that the dynamics underlying the portfolio time series w_t is as simple as possible. According to the realization property, this can be accomplished by penalizing the rank of the matrix in (13). The nonconvexity of the rank function, however, implies that practical algorithms could only be guaranteed to deliver local solutions of a rank-penalized problem. A possible way to deal with this issue is by resorting to a convex relaxation.

B. Nuclear Norm as a Proxy for the Rank Function

For a generic matrix $A \in \mathbb{R}^{D_1\times D_2}$ with SVD (2), the *nuclear norm* (a.k.a. trace norm or Schatten-1 norm) is defined as:

$$\|A\|_* = \sum_{r\in\mathbb{N}_R} \sigma_r \quad (14)$$

and is a convex function of A. The nuclear norm has been used to devise convex relaxations to many optimization problems involving the rank function [19], [24]. This parallels the use of the l_1-norm for cardinality minimization and sparse signal estimation. It can be shown that $\|A\|_*$ is the *convex envelope*[5] on the unit ball of the non-convex function $\mathrm{rank}(A)$ [10]. In the present context, this makes

$$\|\mathcal{H}_R(W)U_H^\perp\|_* \quad (15)$$

a suitable replacement for $\mathrm{rank}(\mathcal{H}_R(W)U_H^\perp)$. Note that $\mathcal{H}_R(W)U_H^\perp$ is a linear function of W; since the composition of a convex function with a linear function is convex, (15) is a convex function.

C. Main Problem Formulation

Our approach consists of solving the optimization problem:

$$\begin{aligned} \max_{W\in\mathbb{R}^{N\times T}} & \left(w_T^\top \hat{m} + \tfrac{1}{T-1}\textstyle\sum_{t\in\mathbb{N}_{T-1}} w_t^\top z_{t+1}\right) - \lambda\|\mathcal{H}_R(W)U_H^\perp\|_* \\ \text{subject to } & w_T^\top \hat{m} + \Phi^{-1}(\beta)\|\hat{\Sigma}^{1/2}w_T\| \geq \alpha \\ & w_{nt} \geq 0, \; n \in \mathbb{N}_N \text{ and } t \in \mathbb{N}_T \\ & 1^\top w_t = 1, \; t \in \mathbb{N}_T \end{aligned} \quad \text{(DynOpt)}$$

where $\lambda \geq 0$, α and $0 < \beta \leq 1/2$ are user-defined parameters. In this problem, $(z_t)_{t\in\mathbb{N}_{T-1}}$ are the past net returns defined in (3), $U_H^\perp$ is the matrix (12) constructed upon inputs $(u_t)_{t\in\mathbb{N}_T}$ and, finally, $\hat{m}$ and $\hat{\Sigma}$ are the market parameters conveying expectation on the market in $T+1$. Notice that the maximization problem (DynOpt) can be reformulated as the problem of minimizing the sum of convex functions subject to convex constraints; therefore, it is equivalent to a convex problem.

The first distinguishable difference with respect to (MkwOpt) and (VaROpt), is that in (DynOpt) we optimize also past portfolios. In this respect note that for $\lambda = 0$, the objective function of (DynOpt) is *separable*; in other words, the optimization over the present and over the past portfolios are mathematically independent problems. In this situation, if $\hat{W} = [\hat{w}_1, \; \hat{w}_2, \; \ldots, \hat{w}_T]$ denotes a solution to (DynOpt), then $\hat{w}_T$ is a solution to (VaROpt) provided that the same value of $\hat{m}$, $\hat{\Sigma}$, α and β are used in the latter.

[5] I.e., the largest convex underestimator.

D. Dynamic Nature of Portfolio Management

For $\lambda > 0$ past and present become entangled; $\|\mathcal{H}_R(W)U_H^\perp\|_*$ is used as a proxy for $\mathrm{rank}(\mathcal{H}_R(W)U_H^\perp)$ and imposes that the dynamics of the time series w_t can be explained by a low-order linear time-invariant state-space system with input u_t. Problem (DynOpt) considers a multiperiod scenario in which realized past returns, together with the expected future return, determine the behavior of the underlying dynamics. More specifically, the objective of (DynOpt) requires that both expected and realized returns are maximized. In contrast, problems (MkwOpt) and (VaROpt) are concerned with single-period investments and regard the market parameters and — consequently — the portfolio, as if they were constant over time, additionally ignoring known influencing factors.

Finally, observe that (DynOpt) models the dynamics *implicitly*, i.e., without explicitly estimating the system's parameters[6]; if needed, these parameters can be recovered via simple algebraic steps in a successive stage, as we discuss later. This feature contrasts our approach to alternative techniques that explicitly model the system. This is the case of procedures based on the Kalman filter, see, e.g., [8]. Note that finding the system's parameters, for instance via Expectation Maximization (EM) [23], inevitably leads to non-convex problem formulations. Avoiding explicit modelling is what ultimately ensures the convexity of our problem formulation.

E. Link With System Identification Methods

The first use of (15) as a penalty function is due to [15]. This paper proposes an approach to deal with the identification of a state-space system from input-output measurements. The first step in this process is to find the model order. The approach in [15] starts from solving:

$$\hat{W}_{\text{id}} = \arg \min_{W \in \mathbb{R}^{S \times T}} \|W - W_{\text{mes}}\|^2 + \lambda \|\mathcal{H}_R(W)U_H^\perp\|_* \quad (16)$$

where W_{mes} is the matrix of measured outputs and $U_H^\perp$ is constructed upon measured inputs. In general, $\mathcal{H}_R(\hat{W}_{\text{id}})U_H^\perp$ is close to rank-deficient and therefore suggests a suitable model order. On the contrary, a classical subspace-based identification procedures [7] works directly with $\mathcal{H}_R(\hat{W}_{\text{mes}})U_H^\perp$, which often does not share such a favorable property.

Note that (16) assumes knowledge of the output measurements of some existing dynamical system. In contrast, (DynOpt) does not work with output measurements; it rather treats the optimization variable as if they were output of an unobserved dynamical system. Nevertheless, once a solution to (DynOpt) has been found, one can estimate the system's parameters for the sake of an a-posteriori analysis of the result. If $\hat{W}$ is a solution to (DynOpt), this can be accomplished by computing the truncated SVD of $\mathcal{H}_R(\hat{W})U_H^\perp$. The system's parameters can then be obtained by simple algebraic steps, see, e.g., [26, Chapter 2.4].

[6]I.e., the system matrices A, B, C, D and the initial state x_1.

F. Implementation Aspects

1) Market Parameters: The optimization problems in (MkwOpt), (VaROpt) and (DynOpt) require to estimate the market parameters $\hat{m}$ and $\hat{\Sigma}$; this task is hindered by the nonstationary of financial data. We considered exponentially-weighted moving average (EWMA) estimators that are simple to implement and partially amend the pitfalls or vanilla estimators [7] [1]. At each time period we compute a new estimate based upon the newly observed data and the past estimate. That is, for each time period we compute:

$$\hat{m}^{(t+1)} = \mu z_t + (1-\mu)\hat{m}^{(t)} \quad (17)$$

$$\hat{\sigma}_{ij}^{(t+1)} = \mu z_{it} z_{jt} + (1-\mu)\hat{\sigma}_{ij}^{(t)} \quad (18)$$

where we took $\mu = 0.2$ and let $\hat{m}^{(1)}$ and $\hat{\Sigma}^{(1)}$ be equal to the Maximum Likelihood (ML) estimators.

2) Computing Solutions: Problem (DynOpt) can be restated as an SDP so that existing primal-dual interior point solvers can be used to compute solutions. One approach is to impose the Hankel structure of $\mathcal{H}_R(W)$ via linear equality constraints and by restating the nuclear norm via an LMI. This can be done by relying on a known result concerning the nuclear norm, see [11, Lemma 2] or [27, Section 3.1]. This approach, however, is inefficient as it requires a large number of auxiliary variables. The paper [15] shows that for problems like (16) one can exploit additional structure in the semidefinite programming formulation which leads to more efficient implementations of interior-point methods. Similar ideas could be applied to (DynOpt). Another way to solve (DynOpt) is by using the Alternating Direction Method of Multipliers (ADMM), see [3] for a recent review. This method uses only first order information and it is therefore preferable for large scale problems.

Finally, being a convex problem, (DynOpt) can be readily implemented via a Matlab-based modeling system for convex optimization such as CVX [6], or YALMIP [16]. In comparison to the aforementioned alternatives, this approach leads to fast prototyping at the price of an increasing computational time. In our experiments we use CVX; the code needed to compute a solution is reported in Appendix.

IV. Case Studies

A. General Setting

The purpose of this section is to discuss the application of DynOpt and to compare it with VaROpt when the same VaR and market parameters are used within the two mathematical programs. Note that it is natural to compare these problems. In fact, as detailed in Section III-C, the two methods output the same present portfolios when $\lambda = 0$.

For each time period we computed the market parameters as described in Section III-F1; in particular, we used 50 observations to compute the initial ML estimators. For VarOpt this fully determines the present portfolio; for DynOpt we

[7]In our experiments we noticed a significant advantage in using EWMA estimators over the standard Maximum Likelihood approach.

additionally considered the past realized returns for 79 periods prior to the present. Results are reported in terms of realized future returns. That is, if T denotes the present, we report the return in $T+1$ according to the portfolio w_T found either by DynOpt or by VaROpt.

We show results only for two representative values of the parameter λ in DynOpt. In our experiments, however, we computed solutions corresponding to a battery of logarithmically distributed parameter values. In line with previous evidence on sparsity-inducing penalties, phase transitions seem to occur. That is, in our experiments we observed that results remain almost unchanged for large spectra of parameter values and then change significantly at given values.

B. Artificial Market Data

1) Generating Model for the Returns: We generated synthetic returns of 5 assets according to the model:

$$z_{t+1} = m + \Omega\, z_t + 0.01\, \Psi\, u_t + 0.01\, \epsilon_t \qquad (19)$$

where u_t, ϵ_t are, respectively, $2/5-$dimensional zero-mean gaussian white noises and we let:

$$\Psi = \begin{bmatrix} 2.18 & -1.52 \\ 2.17 & 0.12 \\ 2.17 & 1.56 \\ -0.65 & 1.2 \\ 0.80 & 3.63 \end{bmatrix}, \; m = \begin{bmatrix} 0.005 \\ 0.005 \\ 0.005 \\ -0.004 \\ -0.01 \end{bmatrix},$$

$$\Omega = \begin{bmatrix} 0.28 & -0.25 & 0.02 & 0.49 & -0.31 \\ -0.24 & 0.08 & 0.34 & -0.34 & 0 \\ 0.24 & 0.16 & 0.09 & -0.07 & -0.41 \\ 0.24 & 0.23 & -0.21 & 0.33 & -0.34 \\ -0.27 & -0.22 & 0.26 & 0.04 & 0 \end{bmatrix}.$$

In (19), the series ϵ_t represents idiosyncratic returns; the process u_t is assumed observed and used as input in DynOpt, together with the previous level of returns.

2) Comparison of Different Strategies: We also considered the situation in which DynOpt is fed with a $7-$dimensional gaussian white noise generated independently from z_t; we let the mean be 0 and set the variance to 0.1; we denote this case by DynOptRnd. Within DynOpt, DynOptRnd and VaROpt we took $\alpha = -0.05$ and $\beta = 0.15$. For all the strategies we included a riskless security which we assumed to have return 0. Figure 2 reports the evolution of wealth of the different approaches. We also include a baseline strategy, denoted by Unif, that works by allocating wealth uniformly over the prescribed assets. Table I reports the *ex-post Sharpe ratios* [22] computed with respect to the returns produced by Unif, used as benchmark. We report results for two different values of the regularization parameter λ in DynOpt. DynOpt with $\lambda = 1$ (thick purple line in Figure 2) significantly outperforms the alternative strategies.

3) Random Inputs as Initial Estimate: In the present paper, we assumed that the input process were observed. This is especially meaningful if one knowns factors predicting asset returns, as in our model (19). Alternatively, one might treat the inputs as unknowns at the price of sacrificing convexity. In the context of Gaussian state-space models [20], for instance, one uses the EM algorithm to alternatively learn the model's parameters and estimate the state and innovation sequences. In this sense, one might regard the independently generated random inputs above as an initial estimate; after having found the model's parameter via algebraic steps, as discussed in Section III-E, one might proceed by finding new input sequences (the innovations) conditioned on the current values of the parameters.

TABLE I: Ex-post Sharpe ratios for the artificial data

VaROpt	DynOpt		DynOptRnd	
	λ			
	0.01	1	0.01	1
0.09	0.27	**0.37**	0.19	0.19

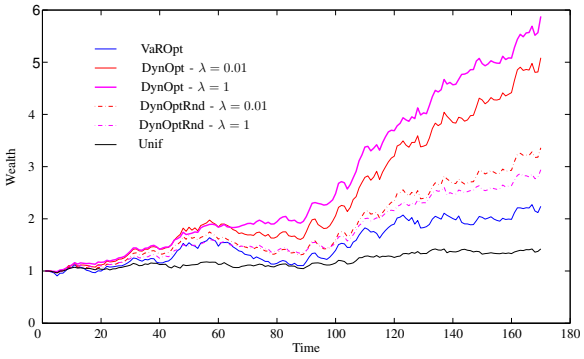


Fig. 1: Evolution of wealth for the different strategies applied on the the artificial data

C. Real Market Data

1) Portfolio and Inputs: In this case we considered a portfolio of 13 ETFs that track indices in different sectors[8]. While there seems to be mixed evidence on the predictability of returns, some papers have found that macroeconomic variables might be relevant within certain asset classes . It has been argued that this might especially be the case for ETFs that are proxies for aggregate of stocks [25]. Here we took the following US macroeconomic variables as inputs for the dynamics of portfolios[9]: 1) GDP, 2) unemployment rate, 3) total consumer credit, 4) industrial production index, 5) consumer price index. We denote this case as DynOptMacro. Alternatively, we considered as input the lagged price values of all the 13 ETFs. We denote this case as DynOptAR. As before, for all the strategies we included a riskless security which we assumed to have return 0, for simplicity.

2) Comparison of Different Strategies: We compared the different methods on a time span that goes from July 2006 to

[8]Yahoo Tickers: DIA, HHH, MDY, QQQ, SPY, XLB, XLE, XLF, XLI, XLK, XLP, XLU, XLY.

[9]We took lagged values so to ensure the feasibility of the investment strategy.

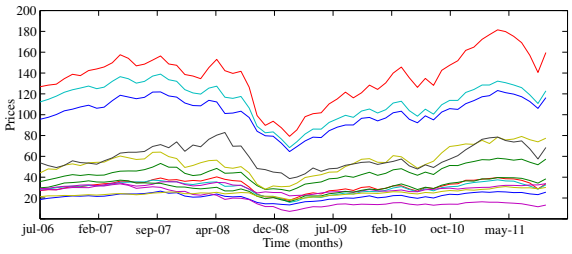


Fig. 2: Prices of ETFs

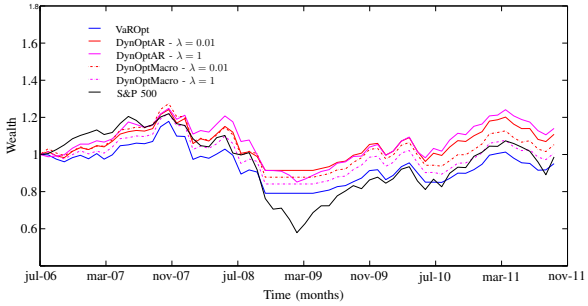
TABLE II: Ex-post Sharpe ratios for the real market data

$\beta = 0.05$

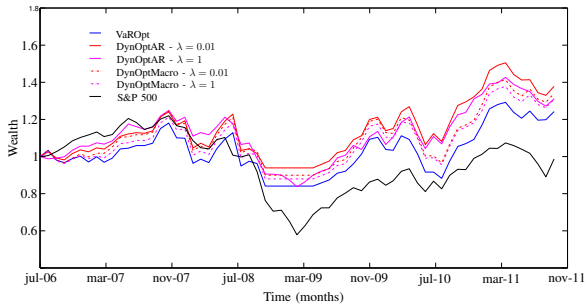
VaROpt	DynOptMacro		DynOptAR	
	λ			
	0.01	1	0.01	1
-0.03	0.01	-0.01	0.03	**0.04**

$\beta = 0.15$

VaROpt	DynOptMacro		DynOptAR	
	λ			
	0.01	1	0.01	1
0.07	0.09	0.09	**0.11**	0.09



(a) $\beta = 0.05$



(b) $\beta = 0.15$

Fig. 3: Evolution of wealth for the different strategies applied on the the real data

November 2011. Within DynOpt and VaROpt we let $\alpha = -.05$ and considered two different risk levels corresponding to $\beta = 0.05$ and $\beta = 0.15$. For DynOpt, 79 months previous to the present month were considered for each time period. Table II reports the ex-post Sharpe ratios computed with respect to the S&P 500. Figure 3 reports the evolution of wealth of the different approaches. Realized future returns of DynOpt generally outperform those of VaROpt; the differences, however, are not as significant as before. Autoregressive inputs seem to be more relevant than macroeconomic variables. Expectedly, the more risky approach ($\beta = 0.15$) on the long run leads to more favorable capital growth. Notably, during the crash of 2008-2009 all the allocation strategies (with the only exception of DynOptAR with $\lambda = 1$) concentrate all resources on the riskless asset, see Figure 3. This holds for both the levels of risk considered.

V. Conclusion

Classical mean-value portfolio optimization problems are concerned with single-period investments. They regard the market parameters and the portfolio as if they were constant over time. In contrast, DynOpt considers a multiperiod scenario and keeps into account the portfolio dynamics. Portfolio weights are regarded as outputs of an underlying dynamical system driven by external inputs. It remains an open problem to study the situation where inputs are unknowns instead of being observed. This is especially relevant for all those cases where it is not possible to find explicit factors explaining asset returns.

Acknowledgment

Research supported by Research Council KUL: GOA/10/09 MaNet , PFV/10/002 (OPTEC), several PhD/postdoc & fellow grants; Flemish Government: IOF: IOF/KP/S-CORES4CHEM; FWO: PhD/postdoc grants, projects: G.0320.08 (convex MPC), G.0557.08 (Glycemia2), G.0588.09 (Brain-machine), G.0377.09 (Mechatronics MPC); G.0377.12 (Structured systems) research community (WOG: MLDM); IWT: PhD Grants, projects: Eureka-Flite+, SBO LeCoPro, SBO Climaqs, SBO POM, EUROSTARS SMART, iMinds 2012. Belgian Federal Science Policy Office: IUAP P7/ (DYSCO, Dynamical systems, control and optimization, 2012-2017). EU: ERNSI, FP7-EMBOCON (ICT-248940), FP7-SADCO (MC ITN-264735), ERC ST HIGHWIND (259 166), ERC AdG A-DATADRIVE-B. COST: Action ICO806: IntelliCIS.

References

[1] C. Alexander. Moving average models for volatility and correlation, and covariance matrices. *Handbook of Finance*, 2008.

[2] M.J. Best and R.R. Grauer. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies*, 4(2):315–342, 1991.

[3] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.

[4] S.P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[5] V.K. Chopra and W.T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The journal of portfolio management*, 19(2):6–11, 1993.

[6] Inc. CVX Research. CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://cvxr.com/cvx, September 2012.

[7] B. De Moor, M. Moonen, L. Vandenberghe, and J. Vandewalle. A geometrical approach for the identification of state space models with singular value decomposition. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 2244–2247. IEEE, 1988.

[8] J. DiLellio. A kalman filter control technique in mean-variance portfolio management. *Journal of Economics and Finance*, pages 1–27, 2012.

[9] E.F. Fama and K.R. French. Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84, 2012.

[10] M. Fazel. *Matrix rank minimization with applications*. PhD thesis, Elec. Eng. Dept., Stanford University, 2002.

[11] M. Fazel, H. Hindi, and S.P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the American Control Conference, 2001*, volume 6, pages 4734–4739, 2001.

[12] D. Goldfarb and G. Iyengar. Robust convex quadratically constrained programs. *Mathematical Programming*, 97(3):495–515, 2003.

[13] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

[14] R.A. Haugen and N.L. Baker. Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3):401–439, 1996.

[15] Z. Liu and L. Vandenberghe. Interior-point method for nuclear norm approximation with application to system identification. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1235–1256, 2009.

[16] J. Lofberg. Yalmip: A toolbox for modeling and optimization in matlab. In *IEEE International Symposium on Computer Aided Control Systems Design*, pages 284–289. IEEE, 2004.

[17] I. Markovsky, J.C. Willems, S. Van Huffel, and B. De Moor. *Exact and approximate modeling of linear systems: A behavioral approach*, volume 11. Society for Industrial Mathematics, 2006.

[18] H.M. Markowitz. Portfolio selection. *J. Finance*, 7:77–91, 1952.

[19] B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52:471–501, 2007.

[20] S. Roweis and Z. Ghahramani. A unifying review of linear gaussian models. *Neural computation*, 11(2):305–345, 1999.

[21] W.F. Sharpe. A simplified model for portfolio analysis. *Management science*, 9(2):277–293, 1963.

[22] W.F. Sharpe. The sharpe ratio. *The Journal of Portfolio Management*, 21(1):49—58, 1994.

[23] R.H. Shumway and D.S. Stoffer. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of time series analysis*, 3(4):253–264, 1982.

[24] M. Signoretto and J. A. K. Suykens. Convex estimation of cointegrated VAR models by a nuclear norm penalty. In *16th IFAC Symposium on System Identification*, pages 95—100, 2012.

[25] S. Sridharan, D. Chitturi, and A.A. Rodriguez. A receding horizon control approach to portfolio optimization using a risk-minimax objective for wealth tracking. In *IEEE International Conference on Control Applications (CCA)*, pages 1282–1287, 2011.

[26] P. Van Overschee and B. De Moor. *Subspace Identification for the Linear Systems: Theory–Implementation–Applications*. Kluwer Academic Publishers, 1996.

[27] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.

[28] M. Verhaegen and P. Dewilde. Subspace model identification part 1. the output-error state-space model identification class of algorithms. *International journal of control*, 56(5):1187–1210, 1992.

APPENDIX
CVX CODE FOR DYNOPT

```
% Uperp is the orthogonal complement of the
% input Hankel matrix ;
% m is the vector of expected future returns ;
% SigmaSqrt is the square root of the
% covariance matrix of future returns ;
% returns is the matrix of realized (past) returns ;

cvx_begin

    variable W(N,T);
    variable ZZ(N*R,L)
    variable HH(N*R,T-R+1)
    variable H1(R,T-R+1) hankel;
    variable H2(R,T-R+1) hankel;

                    .
                    .
                    .

    variable c(T-1);
    minimize(-m'*W(:,end)-mean(c)+nucpar*(norm_nuc(ZZ)))
    subject to

    for n=1:N
        switch n
            case 1
                [H1(1:end,1);H1(end,2:end)']==W(1,:)';
                HH((n-1)+1:N:end,:)==H1;
            case 2
                [H2(1:end,1);H2(end,2:end)']==W(2,:)';
                HH((n-1)+1:N:end,:)==H2;


                    .
                    .
                    .

        end
    end
    ZZ==HH*Uperp;
    for t=1:T-1
        c(t)==returns(:,t+1)'*W(:,t);
    end
     W(:)>=0;
    m'*W(:,end)+norminv(beta)*...
              norm(SigmaSqrt*W(:,end),'fro')>=alpha;
    sum(W,1)==ones(1,T);

cvx_end
```