数理解析研究所講究録
第 1736 巻 2011 年 5-17

Dynamic Investment Strategy for Factor Portfolios with Regime Switches

筑波大学・ビジネス科学研究科 小松 高広 (Takahiro Komatsu)
Graduate School of Business Sciences
University of Tsukuba
筑波大学・ビジネス科学研究科 牧本 直樹 (Naoki Makimoto) *
Graduate School of Business Sciences
University of Tsukuba

Abstract

This paper derives semi-analytic solutions for dynamic optimization of factor portfolios in a mean-variance framework with transaction costs and regime switches which drive discontinuous changes of model parameters. The optimal portfolio is composed of a linear combination of "current" and "target" portfolios the latter of which is more influenced by regimes. For some special cases, we derive analytic solutions that have much simplified form and easy to understand. Numerical experiments are also conducted to confirm economically intuitive sensitivities of the optimal solutions to changes in key parameters that are regime-dependent.

Key Word: Factor portfolio, optimal investment, regime switch.

1 Introduction

Over the course of past decades, the financial markets have exhibited drastic changes in return generating processes that deviate from those in long-term expectations. For example, in late 1970s and early 80s that are known as a lost decade, equity markets were stuck under the stagflation macro economy. In late 1990s, the market participants experienced instability in currencies driven by fragile underpinning in economies across emerging countries. A recent decade includes the US equity market having dropped significantly throughout internet bubble and the global crisis in economy and in the financial markets triggered by subprime loan that turned out to bring out bankruptcy of the Lehman Brothers.

In decision making processes such as asset allocation in both of strategic and tactical investment horizon, investors try to predict returns and estimate risks and transaction costs. Contrary to the drastic and discontinuous behavior in financial markets, the traditional practices in investment management have relied on rather simple models mainly because of their simplicity. Prediction models often consist of a single set of key financial variables such as expected returns, volatility and correlation between assets, or even in dynamic models, key parameters are fixed and financial variables changes continuously.

*This research was supported in part by the Grant-in-Aid for Scientific Research (No. 20510126) of the Japan Society for the Promotion of Science in 2008–2010.

On the other hand, academic endeavor and empirical analyses among several areas in macro economics had already made significant progress in figuring out nature of the drastic and discontinuous changes of economic variables by introducing regime switches. In earlier studies in finance, regime switching models have been applied to wide ranges of assets and markets to successfully explain their dynamic behavior. Figure 1 shows estimated regime probabilities of a 2 regime Markov switching vector autoregressive (VAR) process of 3-factor model by Fama and French [10] and Carhart [7]. We observe jagged regime probabilities flipping between 0 and 1 and some persistent periods to stay either regime. In this example, regime 1 exhibits higher returns, lower volatilities, lower correlation and lower decay rates across three factors while regime 2 shows opposite characteristics.

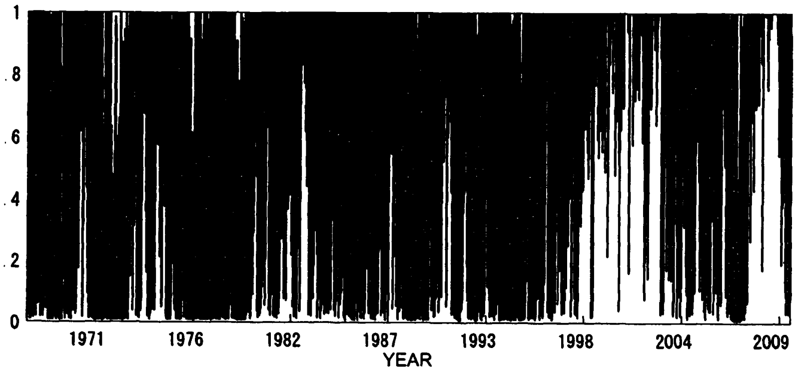

Figure 1: Estimated regime probabilities of 3-factor model by Fama and French [10]

Initiated by Baum and Petrie [5] and extensively studied in the statistics and econometrics literature, e.g., Titterington, Smith and Markov [16] and Hamilton [14], Markov mixture of dynamic models have attracted increasing interest. The model has an advantageous nature of flexibility to approximate a broad range of dynamics in the real world. Ang and Bekaert [1, 2] construct and numerically solve a regime switching model of international equity markets and report that ignoring the regimes could cost under a presence of cash in asset allocation problems. Among well known factors in individual stock markets, market risk, value, small cap and momentum, Arshanapalli, Fobbozi and Nelson [4] reports that the behavior of these premia under different macro economic scenarios is different across factors. The study implies potential presence of different mechanism to drive the equity factor returns from those handled in traditional linear models. Coggi and Manescu [8] presents a state-dependent version of Fama and French [10] model to overcome the shortcoming that the original model exhibits quite poor performance in some periods. Ang and Kristensen [3] finds that time dependency of alpha and beta in the Fama-French model by introducing kernel regression for non-parametric estimation. Although much is left for further research to understand what leads the alpha to deviate from zero, this suggests that some dynamics drive the alpha and beta over time.

Choices of underlying models in the regime switching process are also important decisions to approximate highly complicated actual returns. For example, recently portfolio managers tend to pay attentions not only to cross-sectional information across assets in terms of return forecast but also to time series nature of forecasted returns for measuring persistency of the forecasts. Grinold [13] points out that vintage of information can be found in portfolios under a presence of transaction costs. Sneddon [15] solves mean-variance optimal portfolio problem with transaction costs and reveals that the optimal portfolio should trade fast decay assets more aggressively than slow decay assets.

As well as explaining behaviors of assets in the markets, how optimally investors should behave is similarly important to understand asset returns under the regime switching structure. This paper addresses an optimal portfolio problem with regime switches. Our model extends Gârleanu and Pedersen [12], who derive a closed form solution for the model with multiple securities and multiple return predictors with different mean-reversion speeds, to regime switching structure that drives key parameters in factor returns, transaction costs and investors risk tolerance.

The optimal portfolio is regime-dependent and consists of a linear combination of "current" and "target" portfolios. Decay speed of the vector autoregressive factor process is influenced by parameters that characterize nature of regime switching. Numerical experiments illustrate intuitive behaviors of the solutions. For example, the higher transaction cost anticipated to a regime to switch into, the slower to change asset weights. The less likely to switch into other regimes, the more rapidly asset weights adjust to those in a regime to switch into. The faster decay in factor returns anticipated to a regime to switch into, the smaller asset weights in a current regime. The more likely to switch into other regimes, the closer asset amount to hold in a current regime to those in a regime to switch into.

This paper contributes to literature in several ways. We obtain semi-analytic solutions of dynamic portfolio optimization problems in a mean-variance framework with transaction costs and regime switches. The model assumes regime switches within key parameters in factor portfolios including those related to expected returns modeled by VAR(1) which is sufficiently general to approximate complicated actual returns observed in the markets as well as covariance, transaction costs, factor loadings and investors' risk tolerance. Sensitivity of the optimal solutions to the key parameters are also demonstrated through numerical examples.

The outline of this paper is as follows. Section 2 describes a discrete-time dynamic portfolio optimization problem with regime switches. In Section 3, we solve the problem by dynamic programming and obtain the optimal portfolios. Some special cases are also discussed. Section 4 exhibits numerical examples to investigate properties of the optimal portfolios. Finally, we conclude the paper in Section 5.

2 Model description

We consider an economy with N securities traded at time $t = 1, 2, \ldots$. The price changes of security i between time t and $t+1$ is $r_i(t+1) = p_i(t+1) - p_i(t)$. We assume that an $N \times 1$ return vector $\boldsymbol{r}(t) = (r_1(t), \ldots, r_N(t))^\top$ is given by

$$\boldsymbol{r}(t+1) = \boldsymbol{\theta}(t) + \boldsymbol{\alpha}(t) + \boldsymbol{u}(t), \tag{2.1}$$

where $\boldsymbol{\theta}(t)$ is the fair return from the CAPM, $\boldsymbol{\alpha}(t)$ is the predictable excess return known to the investor at time t and $\boldsymbol{u}(t)$ is an unpredictable zero-mean noise ($\top$ denotes transpose). The predictable excess return $\boldsymbol{\alpha}(t)$ is given by

$$\boldsymbol{\alpha}(t) = \boldsymbol{L}_i \boldsymbol{f}(t), \tag{2.2}$$

where $\boldsymbol{f}(t)$ is an $M \times 1$ vector of factors that predict returns and $\boldsymbol{L}_i$ is an $N \times M$ matrix of factor loadings when the state of the economy (regime) at time t is i. The dynamics of the factor is modeled by

$$\boldsymbol{f}(t+1) = \boldsymbol{\mu}_i + \boldsymbol{\Phi}_i \boldsymbol{f}(t) + \boldsymbol{\epsilon}(t), \tag{2.3}$$

where $\boldsymbol{\mu}_i$ determines the level of the factors, $\boldsymbol{\Phi}_i$ is an $M \times M$ positive-definite matrix of mean-reversion coefficients, and $\boldsymbol{\epsilon}(t)$ is a zero-mean shock affecting the predictors.

As pointed out in Section 1, the financial market sometimes exhibits drastic changes in return generating processes. One useful way to represent such discontinuous behavior in the market is to introduce regime switches. By allowing model parameters being regime-dependent, the dynamics of the model is expected to be well fitted to those of the real market. Following Hamilton [14], the regime process $I(t)$ follows a Markov chain on $\{1, \ldots, J\}$ where the transition probabilities of going from regime i at time t to regime j at time $t+1$ are denoted by $p_{ij} = \mathrm{P}\,(I(t+1) = j \mid I(t) = i)$.

The noise terms $\boldsymbol{u}(t)$ and $\boldsymbol{\epsilon}(t)$ are assumed to be conditionary independent in the sense that, given the regime process $I(s) = i$ and $I(t) = j$, the noise terms $\boldsymbol{u}(s), \boldsymbol{u}(t), \boldsymbol{\epsilon}(s)$ and $\boldsymbol{\epsilon}(t)$ are independent of each other for any s, t, i and j. The covariance matrices of $\boldsymbol{u}(t)$ and $\boldsymbol{\epsilon}(t)$ are, however, regime-dependent and are given by $\boldsymbol{W}_i$ and $\boldsymbol{\Sigma}_i$, respectively, when $I(t) = i$. We also assume that the factor process $\boldsymbol{f}(t)$ is stationary in time. Conditions for the stationarity of Markov-switching vector autoregressive processes are given in Francq and Zakoïan [11].

If $I(t) = i$ and an investor invests $x_i(t)$ to security i at time t, the excess return of the portfolio $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^\top$ between time t and $t+1$ is $\boldsymbol{x}(t)^\top\{\boldsymbol{\alpha}(t) + \boldsymbol{u}(t)\}$, the mean and variance of which are given by

$$\mathrm{E}\left(\boldsymbol{x}(t)^\top\{\boldsymbol{\alpha}(t) + \boldsymbol{u}(t)\}\right) = \boldsymbol{x}(t)^\top \boldsymbol{L}_i \boldsymbol{f}(t), \tag{2.4}$$

$$\mathrm{V}\left(\boldsymbol{x}(t)^\top\{\boldsymbol{\alpha}(t) + \boldsymbol{u}(t)\}\right) = \boldsymbol{x}(t)^\top \boldsymbol{W}_i \boldsymbol{x}(t). \tag{2.5}$$

Trading is costly in the economy and the transaction cost associated with trading $\boldsymbol{x}(t) - \boldsymbol{x}(t-1)$ is given by

$$\frac{1}{2}\{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}^\top \boldsymbol{B}_i \{\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\}, \tag{2.6}$$

where $\boldsymbol{B}_i$ is a positive-definite matrix measuring the level of trading costs. As noted in Gârleanu and Pedersen [12], the trading cost of the form (2.6) is interpreted as a multi-dimensional version of Kyle's λ.

We consider a risk averse investor and let λ_i denote the coefficient of risk aversion in regime i. The investor's objective is to choose a dynamic trading strategy to maximize the present value of all future expected excess returns, penelized for risks and trading costs. Given an initial portfolio $\boldsymbol{x}(0)$, regime $I(1)$, and factor $\boldsymbol{f}(1)$, the objective function to maximize is expressed as

$$\mathrm{E}\Big(\sum_{t=1}^{\infty}\rho^{t-1}\Big[\boldsymbol{x}(t)^\top\boldsymbol{L}_i\boldsymbol{f}(t)-\frac{\lambda_i}{2}\boldsymbol{x}(t)^\top\boldsymbol{W}_i\boldsymbol{x}(t)-\frac{1}{2}\{\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\}^\top\boldsymbol{B}_i\{\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\}\Big]\ \Big|\ \boldsymbol{x}(0),I(1),\boldsymbol{f}(1)\Big),$$

where $\rho\in(0,1)$ is a discount factor. In the next section, we solve the optimization problem under the assumption that the regime is observable. Though the regime of the actual market is not observable, this assumption is not very unrealistic because, as shown in Figure 1, the regime probabilities estimated from the data often exhibit the property that they flips between 0 and 1in the most part of the time interval.

3 Optimal investment strategy

In this section, we obtain the optimal strategy and the value function by solving the Bellman's equation. Some special cases of interest are also investigated where the value function can be much simplified.

3.1 Bellman's equation

Given an initial portfolio $\boldsymbol{y}$, factor $\boldsymbol{f}$, and regime i at time $t=0$, we define the value function by

$$\begin{aligned}V_i(\boldsymbol{y},\boldsymbol{f}) &= \max_{\{\boldsymbol{x}(t)\}}\mathrm{E}\Big(\sum_{t=1}^{\infty}\rho^{t-1}\Big[\boldsymbol{x}(t)^\top\boldsymbol{L}_i\boldsymbol{f}(t)-\frac{1}{2}\boldsymbol{x}(t)^\top\boldsymbol{A}_i\boldsymbol{x}(t)\\ &\quad -\frac{1}{2}\{\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\}^\top\boldsymbol{B}_i\{\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\}\Big]\ \Big|\ \boldsymbol{x}(0)=\boldsymbol{y},I(1)=i,\boldsymbol{f}(1)=\boldsymbol{f}\Big),\end{aligned}\tag{3.1}$$

where $\boldsymbol{A}_i=\lambda_i\boldsymbol{W}_i$. By the principle of optimality, $V_i(\boldsymbol{y},\boldsymbol{f})$ satisfies the Bellman's equation

$$V_i(\boldsymbol{y},\boldsymbol{f})=\max_{\boldsymbol{x}}\Big[\boldsymbol{x}^\top\boldsymbol{L}_i\boldsymbol{f}-\frac{1}{2}\boldsymbol{x}^\top\boldsymbol{A}_i\boldsymbol{x}-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^\top\boldsymbol{B}_i(\boldsymbol{x}-\boldsymbol{y})+\rho\sum_{j=1}^{J}p_{ij}\mathrm{E}\left(V_j(\boldsymbol{x},\boldsymbol{\mu}_i+\boldsymbol{\Phi}_i\boldsymbol{f}+\boldsymbol{\epsilon}_i)\right)\Big]\tag{3.2}$$

where $\boldsymbol{\epsilon}_i$ is a zero-mean noise in (2.3) with the covariance matrix $\boldsymbol{\Sigma}_i$. The guess solution to (3.2) is

$$V_i(\boldsymbol{y},\boldsymbol{f})=-\frac{1}{2}\boldsymbol{y}^\top\boldsymbol{\beta}_i\boldsymbol{y}+\boldsymbol{\delta}_i^\top\boldsymbol{y}+\frac{1}{2}\boldsymbol{f}^\top\boldsymbol{\eta}_i\boldsymbol{f}+\boldsymbol{\xi}_i^\top\boldsymbol{f}+\boldsymbol{y}^\top\boldsymbol{\kappa}_i\boldsymbol{f}+\zeta_i;\qquad i=1,\ldots,J,\tag{3.3}$$

where $\boldsymbol{\beta}_i$ is an $N\times N$ positive definite matrix, $\boldsymbol{\eta}_i$ is an $M\times M$ positive definite matrix, $\boldsymbol{\kappa}_i$ is an $N\times M$ matrix, $\boldsymbol{\delta}_i$ is an $N\times 1$ vector, $\boldsymbol{\xi}_i$ is an $M\times 1$ vector and ζ_i is a scalar. By substituting (3.3) into (3.2) and calculating the expectation with respect to $\boldsymbol{\epsilon}_i$, the right hand side of (3.2) becomes a quadratic function of $\boldsymbol{x}$. The first order optimality condition then yields

$$\boldsymbol{x}_i^*=\rho\boldsymbol{C}_i(\widehat{\boldsymbol{\delta}}_i+\widehat{\boldsymbol{\kappa}}_i\boldsymbol{\mu}_i)+\boldsymbol{B}_i\boldsymbol{y}+(\rho\widehat{\boldsymbol{\kappa}}_i\boldsymbol{\Phi}_i+\boldsymbol{L}_i)\boldsymbol{f};\qquad \boldsymbol{C}_i=(\rho\widehat{\boldsymbol{\beta}}_i+\boldsymbol{A}_i+\boldsymbol{B}_i)^{-1},\tag{3.4}$$

where we define

$$\widehat{z}_i=\sum_{j=1}^{J}p_{ij}z_j;\qquad z=\boldsymbol{\beta}/\boldsymbol{\delta}/\boldsymbol{\eta}/\boldsymbol{\xi}/\boldsymbol{\kappa}/\zeta.\tag{3.5}$$

Substituting (3.4) into the right hand side of (3.2) and equating it to (3.3), we get the following system of equations of unknown coefficient matrices of the value function for $i = 1, \ldots, J$:

$$
\begin{aligned}
\boldsymbol{\beta}_i &= \boldsymbol{B}_i - \boldsymbol{B}_i(\rho\widehat{\boldsymbol{\beta}}_i + \boldsymbol{A}_i + \boldsymbol{B}_i)^{-1}\boldsymbol{B}_i, && (3.6)\\
\boldsymbol{\delta}_i &= \rho\boldsymbol{B}_i\boldsymbol{C}_i(\widehat{\boldsymbol{\delta}}_i + \widehat{\boldsymbol{\kappa}}_i\boldsymbol{\mu}_i), && (3.7)\\
\boldsymbol{\eta}_i &= \rho\boldsymbol{\Phi}_i^\top\widehat{\boldsymbol{\eta}}_i\boldsymbol{\Phi}_i + (\rho\widehat{\boldsymbol{\kappa}}_i\boldsymbol{\Phi}_i + \boldsymbol{L}_i)^\top\boldsymbol{C}_i(\rho\widehat{\boldsymbol{\kappa}}_i\boldsymbol{\Phi}_i + \boldsymbol{L}_i), && (3.8)\\
\boldsymbol{\xi}_i &= \rho\left[\boldsymbol{\Phi}_i^\top(\widehat{\boldsymbol{\eta}}_i^\top\boldsymbol{\mu}_i + \widehat{\boldsymbol{\xi}}_i) + (\rho\widehat{\boldsymbol{\kappa}}_i\boldsymbol{\Phi}_i + \boldsymbol{L}_i)^\top\boldsymbol{C}_i(\widehat{\boldsymbol{\delta}}_i + \widehat{\boldsymbol{\kappa}}_i\boldsymbol{\mu}_i)\right], && (3.9)\\
\boldsymbol{\kappa}_i &= \boldsymbol{B}_i\boldsymbol{C}_i(\rho\widehat{\boldsymbol{\kappa}}_i\boldsymbol{\Phi}_i + \boldsymbol{L}_i), && (3.10)\\
\zeta_i &= \rho\left[\frac{1}{2}\boldsymbol{\mu}_i^\top\widehat{\boldsymbol{\eta}}_i\boldsymbol{\mu}_i + \frac{1}{2}\mathrm{E}\left(\boldsymbol{\epsilon}_i^\top\widehat{\boldsymbol{\eta}}_i\boldsymbol{\epsilon}_i\right) + \widehat{\boldsymbol{\xi}}_i^\top\boldsymbol{\mu}_i + \widehat{\zeta}_i + \frac{1}{2}\rho(\widehat{\boldsymbol{\delta}}_i + \widehat{\boldsymbol{\kappa}}_i\boldsymbol{\mu}_i)^\top\boldsymbol{C}_i(\widehat{\boldsymbol{\delta}}_i + \widehat{\boldsymbol{\kappa}}_i\boldsymbol{\mu}_i)\right]. && (3.11)
\end{aligned}
$$

Hence, the problem of obtaining the optimal portfolio is reduced to solving nonlinear simultaneous equations (3.6) through (3.11).

3.2 Optimal portfolio

From (3.6) to (3.11), the equation we should solve first is (3.6) because all other equations contain $\boldsymbol{\beta}_i$ in $\boldsymbol{C}_i$. Unfortunately, however, (3.6) constitutes a system of nonlinear equations that makes it difficult to obtain an explicit solution. We therefore develop the following procedure for numerically computing $\boldsymbol{\beta}_i$.

Substituting $\boldsymbol{H}_i = \boldsymbol{B}_i^{-1/2}\boldsymbol{\beta}_i\boldsymbol{B}_i^{-1/2}$ and $\boldsymbol{K}_i = \boldsymbol{B}_i^{-1/2}\boldsymbol{A}_i\boldsymbol{B}_i^{-1/2}$ into (3.6), we get after some algebras a system of matrix quadratic equations

$$\rho\boldsymbol{H}_i\widehat{\boldsymbol{H}}_i + \boldsymbol{H}_i(\boldsymbol{K}_i + \boldsymbol{I}) - \rho\widehat{\boldsymbol{H}}_i - \boldsymbol{K}_i = \boldsymbol{O}, \tag{3.12}$$

where we define $\widehat{\boldsymbol{H}}_i = \sum_{j=1}^J p_{ij}\boldsymbol{H}_j$ as before. By solving (3.12), $\boldsymbol{H}_i$ can be expressed in terms of $\boldsymbol{H}_k$ for $k \neq i$ as

$$\boldsymbol{H}_i = -\frac{1}{2\rho p_{ii}}\left[\boldsymbol{D}_i + (1-\rho p_{ii})\boldsymbol{I} \pm \left\{[\boldsymbol{D}_i + (1-a_i)\boldsymbol{I}]^2 + 4a_i\boldsymbol{D}_i\right\}^{1/2}\right], \tag{3.13}$$

where

$$\boldsymbol{D}_i = \boldsymbol{K}_i + \rho\sum_{k=1,k\neq i}^{J} p_{ik}\boldsymbol{H}_k. \tag{3.14}$$

Combining (3.13) and (3.14) leads us to the following iterative procedure for computing $\boldsymbol{H}_i$.

1. Set $\boldsymbol{H}_i^{(0)} = \boldsymbol{O}$ for $i = 1, \ldots, J$ and set $t = 1$.
2. Compute $\boldsymbol{D}_i^{(t)}$ and $\boldsymbol{H}_i^{(t)}$ for $i = 1, \ldots, J$ from

$$
\begin{aligned}
\boldsymbol{D}_i^{(t)} &= \boldsymbol{K}_i + \rho\sum_{k=1,k\neq i}^{J} p_{ik}\boldsymbol{H}_k^{(t-1)}, && (3.15)\\
\boldsymbol{H}_i^{(t)} &= \frac{1}{2\rho p_{ii}}\left[\left\{(\boldsymbol{D}_i^{(t)} + (1-\rho p_{ii})\boldsymbol{I})^2 + 4\rho p_{ii}\boldsymbol{D}_i^{(t)}\right\}^{1/2} - (\boldsymbol{D}_i^{(t)} + (1-\rho p_{ii})\boldsymbol{I})\right] && (3.16)
\end{aligned}
$$

3. The procedure converges if $\boldsymbol{H}_i^{(t-1)}$ and $\boldsymbol{H}_i^{(t)}$ are close enough for all $i = 1, \ldots, J$. Otherwise, increment t by 1 and goto Step 2.

It can be proved by an inductive argument that both $\boldsymbol{D}_i^{(t)}$ and $\boldsymbol{H}_i^{(t)}$ are positive definite that together with the fact that a positive definite matrix has a matrix square root (e.g., Bhatia [6]) assures that $\boldsymbol{H}_i^{(t)}$ in (3.16) is well-defined. Note also that the solution $\boldsymbol{H}_i$ obtained by the procedure is positive definite, and so is $\boldsymbol{\beta}_i = \boldsymbol{B}_i^{1/2}\boldsymbol{H}_i\boldsymbol{B}_i^{1/2}$.

Once $\boldsymbol{\beta}_i$'s are at hand, we can solve (3.6) through (3.11) sequentially. For $K \times L$ matrix $\boldsymbol{M} = [m_{ij}]$, we define a $KL \times 1$ vector by

$$\mathrm{vec}(\boldsymbol{M}) = (m_{11}, \dots, m_{K1}, \dots, m_{1L}, \dots, m_{KL})^\top.$$

First, we solve (3.9) to get

$$\begin{bmatrix} \mathrm{vec}(\boldsymbol{\kappa}_1) \\ \vdots \\ \mathrm{vec}(\boldsymbol{\kappa}_J) \end{bmatrix} = [\boldsymbol{I}_{MNJ} - \rho\boldsymbol{\Gamma}(\boldsymbol{P} \otimes \boldsymbol{I}_{MN})]^{-1} \begin{bmatrix} \mathrm{vec}(\boldsymbol{B}_1\boldsymbol{C}_1\boldsymbol{L}_1) \\ \vdots \\ \mathrm{vec}(\boldsymbol{B}_J\boldsymbol{C}_J\boldsymbol{L}_J) \end{bmatrix}, \tag{3.17}$$

where $\boldsymbol{I}_k$ is an identity matrix of dimension k, $\otimes$ denotes the Kronecker's product, and

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\Phi}_1^\top \otimes (\boldsymbol{B}_1\boldsymbol{C}_1) & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \boldsymbol{\Phi}_J^\top \otimes (\boldsymbol{B}_J\boldsymbol{C}_J) \end{bmatrix}.$$

By solving (3.6), we obtain $\boldsymbol{\delta}_i$'s as

$$\begin{bmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_J \end{bmatrix} = \rho\,[\boldsymbol{I}_{NJ} - \rho\boldsymbol{\Theta}(\boldsymbol{P} \otimes \boldsymbol{I}_N)]^{-1} \begin{bmatrix} \boldsymbol{B}_1\boldsymbol{C}_1\widehat{\boldsymbol{\kappa}}_1\boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{B}_J\boldsymbol{C}_J\widehat{\boldsymbol{\kappa}}_J\boldsymbol{\mu}_J \end{bmatrix}, \tag{3.18}$$

where

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{B}_1\boldsymbol{C}_1 & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \boldsymbol{B}_J\boldsymbol{C}_J \end{bmatrix}.$$

$\boldsymbol{\eta}_i$'s are obtained from (3.7) as

$$\begin{bmatrix} \mathrm{vec}(\boldsymbol{\eta}_1) \\ \vdots \\ \mathrm{vec}(\boldsymbol{\eta}_J) \end{bmatrix} = [\boldsymbol{I}_{M^2J} - \rho\boldsymbol{\Psi}(\boldsymbol{P} \otimes \boldsymbol{I}_{M^2})]^{-1} \begin{bmatrix} \mathrm{vec}((\rho\widehat{\boldsymbol{\kappa}}_1\boldsymbol{\Phi}_1 + \boldsymbol{L}_1)^\top\boldsymbol{C}_1(\rho\widehat{\boldsymbol{\kappa}}_1\boldsymbol{\Phi}_1 + \boldsymbol{L}_1)) \\ \vdots \\ \mathrm{vec}((\rho\widehat{\boldsymbol{\kappa}}_J\boldsymbol{\Phi}_J + \boldsymbol{L}_J)^\top\boldsymbol{C}_J(\rho\widehat{\boldsymbol{\kappa}}_J\boldsymbol{\Phi}_J + \boldsymbol{L}_J)) \end{bmatrix}, \tag{3.19}$$

where

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Phi}_1^\top \otimes \boldsymbol{\Phi}_1^\top & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \boldsymbol{\Phi}_J^\top \otimes \boldsymbol{\Phi}_J^\top \end{bmatrix}.$$

From (3.8), $\boldsymbol{\xi}_i$'s are given by

$$\begin{bmatrix} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_J \end{bmatrix} = \rho\left[\boldsymbol{I}_{MJ} - \rho\boldsymbol{\Phi}^\top\{\boldsymbol{P} \otimes \boldsymbol{I}_M\}\right]^{-1} \begin{bmatrix} \boldsymbol{\Phi}_1^\top\widehat{\boldsymbol{\eta}}_1^\top\boldsymbol{\mu}_1 + (\rho\widehat{\boldsymbol{\kappa}}_1\boldsymbol{\Phi}_1 + \boldsymbol{L}_1)^\top\boldsymbol{C}_1(\widehat{\boldsymbol{\delta}}_1 + \widehat{\boldsymbol{\kappa}}_1\boldsymbol{\mu}_1) \\ \vdots \\ \boldsymbol{\Phi}_J^\top\widehat{\boldsymbol{\eta}}_J^\top\boldsymbol{\mu}_J + (\rho\widehat{\boldsymbol{\kappa}}_J\boldsymbol{\Phi}_J + \boldsymbol{L}_J)^\top\boldsymbol{C}_J(\widehat{\boldsymbol{\delta}}_J + \widehat{\boldsymbol{\kappa}}_J\boldsymbol{\mu}_J) \end{bmatrix}, \tag{3.20}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_1 & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \boldsymbol{\Phi}_J \end{bmatrix}.$$

Finally, we get from (3.11)

$$\begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_J \end{bmatrix} = \rho(\boldsymbol{I}_J - \rho \boldsymbol{P})^{-1} \begin{bmatrix} \frac{1}{2}\boldsymbol{\mu}_1^\top \widehat{\boldsymbol{\eta}}_1 \boldsymbol{\mu}_1 + \frac{1}{2}\mathrm{E}\left(\boldsymbol{\epsilon}_1^\top \widehat{\boldsymbol{\eta}}_1 \boldsymbol{\epsilon}_1\right) + \widehat{\boldsymbol{\xi}}_1^\top \boldsymbol{\mu}_1 + \frac{1}{2}\rho(\widehat{\boldsymbol{\delta}}_1 + \widehat{\boldsymbol{\kappa}}_1 \boldsymbol{\mu}_1)^\top \boldsymbol{C}_1 (\widehat{\boldsymbol{\delta}}_1 + \widehat{\boldsymbol{\kappa}}_1 \boldsymbol{\mu}_1) \\ \vdots \\ \frac{1}{2}\boldsymbol{\mu}_J^\top \widehat{\boldsymbol{\eta}}_J \boldsymbol{\mu}_J + \frac{1}{2}\mathrm{E}\left(\boldsymbol{\epsilon}_J^\top \widehat{\boldsymbol{\eta}}_J \boldsymbol{\epsilon}_J\right) + \widehat{\boldsymbol{\xi}}_J^\top \boldsymbol{\mu}_J + \frac{1}{2}\rho(\widehat{\boldsymbol{\delta}}_J + \widehat{\boldsymbol{\kappa}}_J \boldsymbol{\mu}_J)^\top \boldsymbol{C}_J (\widehat{\boldsymbol{\delta}}_J + \widehat{\boldsymbol{\kappa}}_J \boldsymbol{\mu}_J) \end{bmatrix}, \tag{3.21}$$

where $\mathrm{E}\left(\boldsymbol{\epsilon}_i^\top \widehat{\boldsymbol{\eta}}_i \boldsymbol{\epsilon}_i\right) = \sum_{j=1}^M \sum_{k=1}^M (\widehat{\boldsymbol{\eta}}_i)_{jk} (\boldsymbol{\Sigma}_i)_{jk}$.

The optimal portfolio in (3.4) is expressed as

$$\boldsymbol{x}_i^* = (\boldsymbol{I} - \boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i)\boldsymbol{y} + \boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i\{\boldsymbol{\beta}_i^{-1}(\boldsymbol{\kappa}_i \boldsymbol{f} + \boldsymbol{\delta}_i)\}. \tag{3.22}$$

(3.22) implies that the optimal portfolio is the weighted average of the current portfolio $\boldsymbol{y}$ and the target portfolio $\boldsymbol{\beta}_i^{-1}(\boldsymbol{\kappa}_i \boldsymbol{f} + \boldsymbol{\delta}_i)$ with an weight matrix $\boldsymbol{B}_i^{-1}\boldsymbol{\beta}_i$.

3.3 Regime independent cost parameters

When the covariance matrix $\boldsymbol{W}_i$ of a noise $\boldsymbol{u}(t)$ in (2.1), a risk aversion coefficient λ_i and the transaction cost matrix $\boldsymbol{B}_i$ are regime-independent, i.e., $\boldsymbol{W} = \boldsymbol{W}_i$, $\lambda = \lambda_i$, $\boldsymbol{B} = \boldsymbol{B}_i$, we can get the optimal portfolio explicitly. Because $\boldsymbol{B}$ is regime-independent, (3.12) is reduced to a single matrix quadratic equation which has the explicit positive definite solution

$$\boldsymbol{H} = \frac{1}{2\rho}\left[\left\{(\boldsymbol{K} + (1-\rho)\boldsymbol{I})^2 + 4\rho\boldsymbol{K}\right\}^{1/2} - (\boldsymbol{K} + (1-\rho)\boldsymbol{I})\right]. \tag{3.23}$$

Thus, both $\boldsymbol{\beta}$ and $\boldsymbol{C}$ becomes regime-independent and are given by

$$\boldsymbol{\beta} = \boldsymbol{B}^{1/2}\boldsymbol{H}\boldsymbol{B}^{1/2}, \qquad \boldsymbol{C} = (\rho\boldsymbol{\beta} + \boldsymbol{A} + \boldsymbol{B})^{-1}, \tag{3.24}$$

where $\boldsymbol{A} = \lambda\boldsymbol{W}$. Though other coefficients $\boldsymbol{\delta}_i$, $\boldsymbol{\eta}_i$, $\boldsymbol{\xi}_i$, $\boldsymbol{\kappa}_i$, ζ_i are regime-dependent, they can be obtained from (3.17) to (3.21). The optimal portfolio in this case is

$$\boldsymbol{x}_i^* = (\boldsymbol{I} - \boldsymbol{B}^{-1}\boldsymbol{\beta})\boldsymbol{y} + \boldsymbol{B}^{-1}\boldsymbol{\beta}\{\boldsymbol{\beta}^{-1}(\boldsymbol{\kappa}_i \boldsymbol{f} + \boldsymbol{\delta}_i)\}, \tag{3.25}$$

where the weight matrix $\boldsymbol{B}^{-1}\boldsymbol{\beta}$ of the current portfolio $\boldsymbol{y}$ is also regime-independent.

3.4 Transaction cost matrix proportional to the covariance matrix

In addition to the regime-independence assumptions in Section 3.3, we further assume that the transaction cost matrix $\boldsymbol{B}$ is proportional to the covariance matrix $\boldsymbol{W}$. See Gârleanu and Pedersen [12] for the justification of this assumption.

A straightforward calculation shows that matrix $\boldsymbol{\beta}$ and $\boldsymbol{C}$ are given by

$$\boldsymbol{\beta} = \frac{\sqrt{\{\lambda + (1-\rho)\gamma\}^2 + 4\rho\lambda\gamma} - \{\lambda + (1-\rho)\gamma\}}{2\rho}\boldsymbol{W}, \tag{3.26}$$

$$\boldsymbol{C} = \frac{2}{\sqrt{\{\lambda + (1-\rho)\gamma\}^2 + 4\rho\lambda\gamma} + \lambda + (1+\rho)\gamma}\boldsymbol{W}^{-1}. \tag{3.27}$$

The optimal portfolio in this case then becomes

$$\boldsymbol{x}_i^* = \frac{2\gamma}{\sqrt{\{\lambda + (1-\rho)\gamma\}^2 + 4\rho\lambda\gamma} + \lambda + (1+\rho)\gamma}\boldsymbol{y} + \frac{1}{\gamma}\boldsymbol{W}^{-1}(\boldsymbol{\kappa}_i\boldsymbol{f} + \boldsymbol{\delta}_i), \tag{3.28}$$

that states that, independent of the current regime i, it is optimal to hold fixed portion of the current portfolio $\boldsymbol{y}$.

4 Numerical experiments

Because optimal solutions are regime-dependent in general, it is difficult to intuitively understand how model parameters affect the optimal investment behavior. In order to understand the optimal solutions from financial viewpoint, we conduct numerical experiments. As a benchmark of the experiments, we set the base case with 5 assets ($N = 5$), 3 regimes ($I = 3$) and 3 factors ($M = 3$). Other parameters in the base case are summarized in Appendix A.

As we show in Figure 1, the regime switches observed in the real market flips one to the others frequently. For understanding the nature of the optimal solution, let us suppose that the regime switches take place as slowly as the factor process reaches at the stationary mean. Figure 2 shows optimal holding of asset 5 if the process starts at regime 1 and changes into regime 2 at $t = 5$ when the residual terms of factor returns are ignored. In addition to the base case, two other cases are plotted. The dashed line shows the trajectory when the process remains in regime 1 while the dotted line shows trajectories when the process changes to regime 3 at $t = 5$. Because a transaction cost is lowest in regime 1 and highest in regime 3, the higher transaction cost anticipated to regime switches into, the slower to reduce asset weights.

Another aspect of the optimal solutions that is worth to discuss is how transition probabilities influence the asset allocation. We show four trajectories including the one for the base case and three others with different transition probability matrices in Figure 3.

$$\boldsymbol{P}_{base} = \begin{bmatrix} 0.975 & 0.0125 & 0.0125 \\ 0.0250 & 0.950 & 0.0250 \\ 0.0375 & 0.0375 & 0.925 \end{bmatrix}, \quad \boldsymbol{P}_1 = \begin{bmatrix} 0.9990 & 0.0005 & 0.0005 \\ 0.0250 & 0.9500 & 0.0250 \\ 0.0375 & 0.0375 & 0.9250 \end{bmatrix},$$

$$\boldsymbol{P}_2 = \begin{bmatrix} 0.9000 & 0.0500 & 0.0500 \\ 0.0750 & 0.8500 & 0.0750 \\ 0.0875 & 0.0875 & 0.8250 \end{bmatrix}, \quad \boldsymbol{P}_3 = \begin{bmatrix} 0.8000 & 0.1000 & 0.1000 \\ 0.1250 & 0.7500 & 0.1250 \\ 0.1350 & 0.1375 & 0.7250 \end{bmatrix}.$$

$\boldsymbol{P}_1$ indicates the most sticky among others and $\boldsymbol{P}_3$ represents the most transient one. One should note that significant difference is found not only in the speed of convergence of the trajectories but also in the level of initial holding of asset 5. The less likely to switch into other regimes,

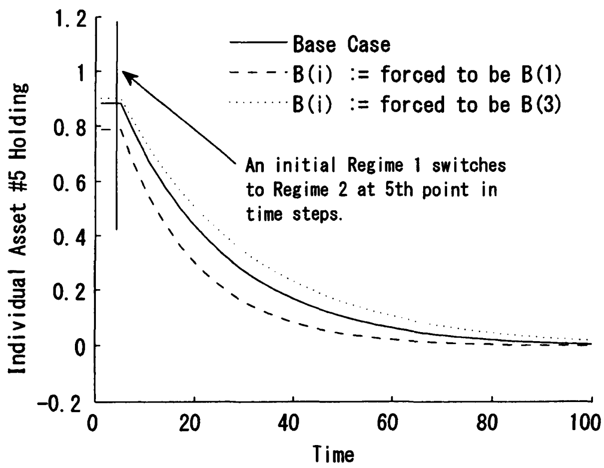

Figure 2: Trajectories of holding of asset 5.

amount of current holding before the regime switch is less affected by other regimes to switch into, and the more rapidly asset weights adjust to those in a switching regime once the switch comes out.

Next we are interested in how return generating processes in factors contribute to form optimal portfolios. In our model, the factor process is modeled by VAR(1) with regime-dependent transition coefficient matrix $\mathbf{\Phi}_i$ that plays important roles in portfolio construction. Figure 4 shows optimal portfolios when all regime-dependent parameters are fixed over time except for $\mathbf{\Phi}_i$. Bars in regime i ($i = 1, 2, 3$) respectively show how optimal holdings of 5 assets should be if $\mathbf{\Phi}_i$ continues all the time. Among others, $\mathbf{\Phi}_1$ has the highest level of autoregressive coefficients while $\mathbf{\Phi}_3$ has the lowest level, i.e., zero autocorrelation. Amount of optimal holding decreases as the level of $\mathbf{\Phi}_i$ decreases from positive to 0. The faster decay in factor returns anticipated to a regime to be switched into, the smaller asset weights in a current regime.

5 Conclusion

Our study incorporates into Gârleanu and Pedersen [12] a regime switching structure and mean-reverting levels to derive a semi-analytic solution in the mean-variance framework with transaction costs. The optimal portfolio is a linear combination of a "current" portfolio and a "target" portfolio which is more influenced by regimes that are more likely to take place than others. Decay speed of the factor process is influenced by parameters that characterize nature of regime switches. Through numerical experiments, we found that the higher transaction cost anticipated to a regime to switch into, the slower to change asset weights. Moreover, the less likely to switch into other regimes, the more rapidly asset weights adjust to those in a regime to switch into.

The contribution of our paper is two folds. First, we derive a semi-analytic solution for optimal portfolios which is easy to compute. From practical viewpoint, our observations about the properties of the optimal portfolio are useful for investment practices because our model is

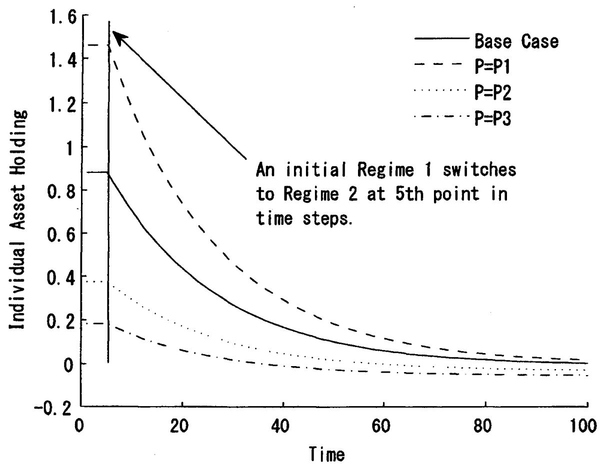

Figure 3: Trajectories of holding of asset 5.

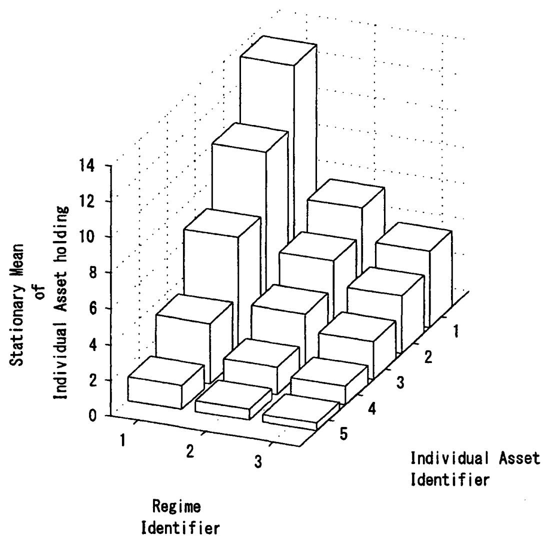

Figure 4: Portfolio selection for different coefficient matrices of the factor return process.

sufficiently flexible to approximate complex behaviors observed in the actual financial markets.

References

[1] Ang, A. and G. Bekaert, "International asset allocation with regime shifts," *Review of Financial Studies*, **15**, 1137–1187, 2002.

[2] Ang, A. and G. Bekaert, "How do regime shifts affect asset allocation?," Working Paper, Columbia University, 2002.

[3] A. Ang and D. Kristensen, "Testing conditional factor models," Working Paper, Columbia University, 2010.

[4] Arshanapalli, B., F.J. Fabozzi and W. Nelson, "The value, size and momentum spread during distressed economic periods," *Finance Research Letters*, **3**, 244–252, 2006.

[5] Baum L.E. and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *The Annals of Mathematical Statistics*, **37**, 1554–1563, 1966.

[6] Bhatia, R., *Matrix Analysis*, Springer, 1997.

[7] Carhart, M.M., "On persistence in mutual fund performance," *Journal of Finance*, **52**, 57–82, 1997.

[8] Coggi P. and B. Manescu, "A multifactor model of stock returns with stochastic regime switching," Working Paper, St. Gallen University, 2004.

[9] Fama E.F. and K.R. French, "The cross-section of expected stock returns," *Journal of Finance*, **47**, 427–465, 1992.

[10] Fama E.F. and K.R. French, "Multifactor explanations of asset pricing anomalies," *Journal of Finance*, **51**, 55–84, 1993.

[11] Francq, C. and J.-M. Zakoïan, "Stationarity of multivariate Markov-switching ARMA models," *Journal of Econometrics*, **102**, 339–364, 2001.

[12] Gârleanu, N. and L.H. Pedersen, "Dynamic trading with predictable returns and transaction costs," *NYU Working Paper No.FIN-09-26*, New York University, 2009.

[13] Grinold, R., "Dynamic portfolio analysis," *Journal of Portfolio Management*, **34**, 12–26, 2007.

[14] Hamilton, J.D., *Time Series Analysis*, Princeton University Press, 1994.

[15] Sneddon, L., "The tortoise and the hare: Portfolio dynamics for active managers," *Journal of Investing*, **17**, 106–111, 2008.

[16] Titterington, D.M., A.F.M. Smith and U.E. Markov, *Statistical Analysis of Finite Mixture Distribution*, Wiley, New York, 1985.

A Parameters in the base case in Section 4

This section summarizes the parameters in the base case in Section 4.

$$
\begin{aligned}
&\boldsymbol{\Phi}_i = \phi_i \boldsymbol{I}; \quad \phi_1 = 0.9, \ \phi_2 = 0.5, \ \phi_3 = 0, \\
&\boldsymbol{\mu}_i = \frac{(1-\phi_i)\mu_i}{250}\mathbf{1}; \quad \mu_1 = 0.25, \ \mu_2 = 0, \ \mu_3 = -0.25, \\
&\boldsymbol{\Sigma}_i = \frac{\sigma_i}{250}\begin{bmatrix} 1 & r_i & r_i \\ r_i & 1 & r_i \\ r_i & r_i & 1 \end{bmatrix}; \\
&\sigma_1 = \frac{0.01(1-\phi_1^2)}{250}, \ \sigma_2 = \frac{0.05(1-\phi_2^2)}{250}, \ \sigma_3 = \frac{0.10(1-\phi_3^2)}{250}, \ r_1 = -0.3, \ r_2 = 0, \ r_3 = 0.3, \\
&\boldsymbol{L}_1 = \boldsymbol{L}_2 = \boldsymbol{L}_3 = \begin{bmatrix} 1 & 1 & 1 \\ .75 & .75 & .75 \\ .50 & .50 & .50 \\ .25 & .25 & .25 \\ .10 & .10 & .10 \end{bmatrix}, \\
&\boldsymbol{W}_i = \boldsymbol{L}_i \boldsymbol{\Sigma}_i \boldsymbol{L}_i^\top + \frac{0.5}{250}\boldsymbol{I}, \\
&\boldsymbol{B}_i = b_i \boldsymbol{I}; \quad b_1 = 0.10, \ b_2 = 0.15, \ b_3 = 0.20, \\
&\boldsymbol{P} = \begin{bmatrix} 0.975 & 0.0125 & 0.0125 \\ 0.0250 & 0.950 & 0.0250 \\ 0.0375 & 0.0375 & 0.925 \end{bmatrix}, \\
&\lambda_1 = 0.10, \ \lambda_2 = 0.05, \ \lambda_3 = 0.01, \\
&\rho = 0.9.
\end{aligned}
$$

Here, $\boldsymbol{I}$ and $\mathbf{1}$ respectively stand for an identity matrix and a column vector of 1's of an appropriate dimension.

For the sake of clarity and simplicity, factor loadings for the individual assets are set to be regime-independent. Some parameters are divided by 250 to transform the annual data to daily ones. In the base case, regime 1 exhibits a tranquil state with higher returns, lower volatilities and correlation, regime 3 exhibits an adverse state and regime 2 inbetween.