Markov Regime Switching Model

Implementation to the Stockholm Stock Market & Comparison with Equal Weight Portfolio

Bachelor Thesis
Sanna Brandel
2017-01-20

Contents

Abstract

The unpredictable behaviour of financial time series has long been a concern for econometricians, making it difficult to find appropriate models with a satisfactory fit. The Markov regime switching model is a popular approach, much in behalf of the way it takes the shifts in the time series behaviour into account.

The model in this thesis is based on a mixture of normal distributions, extended to include a Markov switching behaviour. As the behaviour of the time series changes, regime switches are assigned to it, making the time series alternate between a predetermined number of states.

After the implementation, on two portfolios à seven stocks selected from the Stockholm stock market, the examinations indicated that the fit of the model could be improved by changing the number of states assumed in the estimation. It was found that a Markov regime switching model with three states had the most satisfactory fit to the time series. Subsequently, one of the modelled portfolios was allocated to maximize the Sharpe ratio. This led to some unfavourable extreme allocations, and upon comparison with a portfolio of equal weights containing the same assets the results were poor. Despite a higher yearly return, the modelled portfolio displayed significantly larger volatilities, leaving the results of this evaluation inconclusive. Nevertheless, the implementation lead to a significant improvement in the autocorrelation of the absolute residuals, along with giving the residuals a substantially more homogenic appearance. These results indicate that most of the significant dependence structure has been captured, in particular by the three-state model.

1 Introduction
The Markov regime switching model, first described by G. Lindgren, 1978, is a type of specification in which the main point is handling processes driven by different states, or regimes, of the world. In this model, the observed time series are assumed to follow a non-linear stationary process. The behaviour of the time series is characterized by multiple equations, decided by the different states of the model.

What separates the Markov regime switching model from other switching models is that the switching mechanism is controlled by an unobservable variable that follows a hidden Markov chain. By Markov properties, the current value of the variable depends only on its immediate past value. This means that a structure in the series may prevail for a random period of time, before being replaced by another structure when a switching takes place. This way, the Markov regime switching model is able to capture more complex dynamic patterns.

Financial time series occasionally display dramatic breaks in their behaviour, due to e.g. financial crises. Therefore, the idea of the financial market finding itself in different states at different times becomes appealing. Furthermore, it has been found that financial time series exhibit some formalised facts which can advantageously be reproduced by a hidden Markov model. This has made the Markov regime switching model one of the most popular nonlinear time series models in the literature (Cont, 2001, Hamilton, 1989, 2005, Lindgren, 1978).

1.1 Thesis Statement
The aim with this thesis is to explain the Markov regime switching model in a detailed and comprehensible way, and provide a complete description of the practical implementation to the Stockholm stock market. A portfolio containing seven stocks will be modelled according to a chosen Markov regime switching model. This portfolio will then be allocated to maximize the Sharpe ratio, and finally compared to a portfolio of equal weights.

2 The Hidden Markov Model

A hidden Markov model (HMM) is a bivariate discrete time process $\{S_t, Y_t\}_{t\geq 0}$, where $\{S_t\}$ is an underlying Markov chain and $\{Y_t\}$ is a sequence of independent random variables, of which follows that the conditional distribution of Y_t only depends on S_t. Since the Markov chain S_t is hidden, only the stochastic process $\{Y_t\}$ is available to the observer. In other words, the state of the process is not directly visible, but the output process, dependent on the state, is visible. This means that all statistical inference has to be done in terms of the output stochastic process $\{Y_t\}$ only, as $\{S_t\}$ is not possible to observe (Rydén et al, 2005).

A HMM has an interesting dependence structure, which comes handy when dealing with e.g. financial time series. To get an intuitive hint as to how this dependence works, it is here represented by a graphical model:

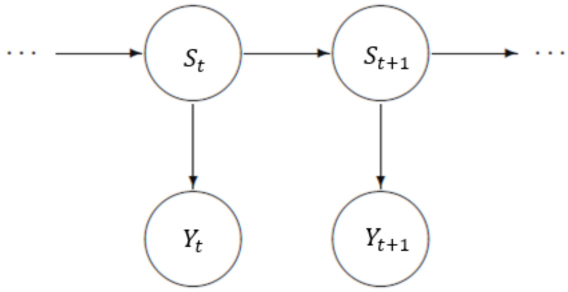

Fig. 2.1 Graphical representation of the dependence structure of a HMM.

As Figure 2.1 implies, the distribution of a variable S_{t+1} conditional on the history of the process S_0, . . . , S_t, is determined only by the value of the preceding variable, S_t. This is all according to the Markov property, where future events are completely independent of the past, depending only on the present state. In addition, the distribution of Y_t conditionally on the past observations Y_0, . . ., Y_{t-1} and the past values of the state, S_0, . . ., S_t, is determined by S_t only (Rydén et al, 2005). Putting this into mathematical terms, we get the following properties:

$$f(S_{t+1}|S_t, \dots, S_1) = f(S_{t+1}|S_t) \quad (1)$$

$$f(Y_t|S_{t-1}, \dots, S_1, Y_{t-1}, \dots, Y_1) = f(Y_t|S_t) \quad (2)$$

2.1 Assumptions of the Hidden Markov Model

A few assumptions on the HMM used in this thesis need to be specified in order to make use of the model.

First of all, the hidden Markov chain is assumed to be time-independent. This means that the transition probabilities of the chain;

$$p_{i,j} = P(S_t = j|S_{t-1} = i) = P(S_t = j|S_{t-1} = i,\, S_{t-2} = k, \dots, S_1 = l\,) \quad (3)$$

between two states i and j in a finite state space Ω = {1, …, N} needs to be constant over time. This is convenient, since said transition probabilities and the Markov chain's set of initial probabilities;

$$\pi_i = P(S_1 = i)$$

are all that is needed to define the dynamic of the HMM.

Secondly, the Markov chain is assumed to be ergodic, i.e. aperiodic and positive recurrent. This is necessary in order to ensure consistency of the estimates of the model (Campigotto, 2009).

3 Definitions

Before postulating a more specific model there are a few definitions to be stated that will be referred to throughout the thesis:

Definition 3.1 A HMM's *filter probability*, in this thesis referred to as $\alpha_{i,t}$, is defined as the probability of the underlying Markov chain being in a given state i at time t, conditionally on a set of observations $\{y_{1:t}\}$; $\alpha_{i,t} = P(S_t = i \mid y_{1:t})$.

Definition 3.2 The *smoothing probability* of a HMM, referred to as $\pi_{i,t}$, determines the probability of the underlying Markov chain being in a given state i at time t, conditionally on a set of observations $\{y_{1:t+\tau}\}$; $\pi_{i,t} = P(S_t = i \mid y_{1:t+\tau}),\ \tau > 0$.

Definition 3.3 The *Bayesian information criterion* (BIC) is a criterion for model selection among a finite set of models, where the model with the lowest BIC is preferred. While adding parameters to a model may increase the likelihood, it may also result in overfitting. BIC solves this issue by introducing a penalty term for the number of parameters. Formal definition: $BIC = -2 \cdot ln(L^{max}) + k \cdot ln(n)$, where $L^{max} =$ the maximized value of the likelihood function of the model, $n =$ the number of observations and $k =$ the number of parameters to be estimated.

4 Method: Specification of Chosen Markov Regime Switching Model

The model applied in this thesis is grounded on a mixture of normal distributions, based mainly on Campigotto, 2009, Hamilton, 2005 and Perlin, 2015.

Assume that the typical historical behaviour of a financial time series can be described by the following process:

$$Y_t = \mu_{S_t} + \varepsilon_t \tag{4}$$

Where;

Y_t is the observed return of the time series at time t

μ_{S_t} is the intercept, or expected return, while in state S_t

ε_t is a normal random stochastic variable, ε_t ~ N(0, $\sigma_{S_t}^2$)

This is a simple case of a model with a switching dynamic. The model in equation (4) is switching states with respect to an indicator value S_t, meaning that with N states there will be N values for μ_{S_t} and $\sigma_{S_t}^2$. Here, the residuals ε_t are assumed to be normal distributed.

4.1 Markov Regime Switching Model with N Regimes

Now, assume that the number of states (or regimes) is N, i.e. S_t ∈ Ω = {1, …, N}. This implies that e.g. the log returns of a financial time series are drawn from N distinct normal distributions, depending on what state the HMM is currently in. This would give us the following model to work with:

$$Y_t = \mu_1 + \varepsilon_t \quad \text{for state 1} \tag{5}$$

$$Y_t = \mu_2 + \varepsilon_t \quad \text{for state 2} \tag{6}$$

$$\vdots$$

$$Y_t = \mu_N + \varepsilon_t \quad \text{for state N} \tag{7}$$

Where;

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_1^2) \quad \text{for state 1} \tag{8}$$

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_2^2) \quad \text{for state 2} \tag{9}$$

$$\vdots$$

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_N^2) \quad \text{for state N} \tag{10}$$

This means that when the state of the HMM for time t is 1, then the expectation of the dependent variable is μ_1 and the variance of the innovations is σ_1^2, etc.

Since the underlying Markov chain is hidden one cannot observe what state the HMM is in directly, but only deduce its operation through the observed behaviour of Y_t. In order to attain the probability law governing the observed data Y_t a probabilistic model of what causes the change from state S_t = i to state S_t = j is required. This can be specified using the transition probabilities of an N state HMM (Hamilton, 2005);

$$p_{i,j} = P(S_t = j | S_{t-1} = i) \quad i, j \in \Omega = \{1, 2, \dots, \mathrm{N}\} \tag{11}$$

The transition probability (11) is by the Markov property described in (3) dependent of the past only through the value of the most recent state. This is one of the central points of the structure of a Markov regime switching model, i.e. the switching of the states of the underlying HMM is a stochastic process itself.

5 Parameter Estimation

There are several ways to estimate the required parameters of the N-state Markov regime-switching model given by (5) - (10), e.g. by using the EM algorithm from Dempster et al., 1977. In the framework of this thesis, however, the parameters will be estimated using maximum likelihood.

5.1 Maximum Likelihood Estimation

Consider the model given by equation (5) – (10), i.e. a Markov regime-switching model with N regimes. The estimation will be performed using Hamilton's filter, where the main idea is to calculate each state's filter probabilities by making inferences on each state's unknown probabilities based on the available information. When the filter probabilities are obtained, we have the probabilities one needs for calculating the log likelihood of the model.

5.1.1 Calculating the Filter Probabilities

The model's filter probabilities, defined in Definition 3.1, are calculated by utilizing the model's iterative relations by means of recursion. This can be done using a combination of the relation between observations and hidden states, and the endogenous relation between hidden states, demonstrated in equation (1) and (2) respectively.

Begin from the starting value in our recursion, i.e. with the probability of being in state i at time t=1:

$$\begin{aligned}\alpha_{i,1} &= P(S_1 = i \mid y_1) \\ &= \frac{f(\, S_1 = i,\ y_1)}{f(y_1)} \\ &= \frac{f(y_1 \mid S_1 = i\,)P(S_1 = i)}{\sum_{j=1}^{N} f(S_1 = i,\ y_1)} \\ &= \frac{f(y_1 \mid S_1 = i\,)P(S_1 = i)}{\sum_{j=1}^{N} f(y_1 \mid S_1 = j\,)P(S_1 = j)}\end{aligned}$$

The second element of the numerator is simply the previously mentioned initial probability of the Markov chain, i.e. $P(S_1 = i) = \pi_i$, and it will henceforth be denoted as such. One can at this point notice that $\alpha_{i,1}$ is the normalized value of the product between the initial probability and the conditional probability function $f(Y_1 \mid S_1 = i\,)$, and can therefore be written as follows:

$$\alpha_{i,1} = \frac{f(Y_1 \mid S_1 = i\,)\pi_i}{\sum_{j=1}^{N} f(Y_1 \mid S_1 = j\,)\pi_j} = \lceil f(Y_1 \mid S_1 = i\,)\pi_i \rceil$$

Now, assume that we know the filter probability at time t-1, namely $\alpha_{i,t-1}$. Following the same strategy as for t = 1 leads to the following recursion:

$$\begin{aligned}\alpha_{i,t} = P(S_t = i \mid y_{1:t}) &= \frac{f(S_t = i, y_t \mid y_{1:t-1})}{f(y_t \mid y_{1:t-1})} \\ &= \frac{f(S_t = i, y_t \mid y_{1:t-1})}{\sum_{j=1}^{N} f(S_t = j, y_t \mid y_{1:t-1})} \\ &= \lceil f(S_t = i, y_t \mid y_{1:t-1}) \rceil\end{aligned}$$

$$= [f(y_t \mid S_t = i, y_{1:t-1})P(S_t = i \mid y_{1:t-1})]$$

$$= [f(y_t \mid S_t = i)P(S_t = i \mid y_{1:t-1})] \quad (12)$$

By the Markov property in equation (2), also demonstrated in Fig. 2.1, one can deduce that once observation y_t has been extracted, the only relation between the current state of the hidden chain S_t, and the set of observations, $\{y_{1:t-1}\}$, exists through S_{t-1}. Following this understanding, the next step is to choose $\alpha_{i,t-1}$:

$$\alpha_{i,t} = P(S_t = i \mid y_{1:t-1}) =$$

$$= \sum_{j=1}^{N} P(S_t = i, S_{t-1} = j \mid y_{1:t-1})$$

$$= \sum_{j=1}^{N} P(S_t = i \mid S_{t-1} = j,\ y_{1:t-1})P(S_{t-1} = j \mid y_{1:t-1})$$

$$= \sum_{j=1}^{N} P(S_t = i \mid S_{t-1} = j)P(S_{t-1} = j \mid y_{1:t-1})$$

Here, the first part of the sum corresponds to the transition probability $p_{i,j}$ (11) between state i and j, and the second part to the filter probability at time t-1, $\alpha_{i,t-1}$. Inserting this into formula (12) the recursion relation becomes (Campigotto, 2009):

$$\alpha_{i,t} = [f(y_t \mid S_t = i\,) \textstyle\sum_{j=1}^{N} p_{i,j}\alpha_{i,t-1}]$$

5.1.2 The Maximum Likelihood Formula

We denote the set of parameters of the N-state Markov regime-switching model given by (5) - (10) we would like to estimate by {Θ}. In our case, {Θ} = $\{(\pi_i), (p_{i,j}), (\mu_i^m), (\Sigma_i^{m,n})\}$, i, j ∈ $\Omega = \{1, 2, \dots, \mathrm{N}\}, \mathrm{m}, \mathrm{n} \in \{1, 2, \dots, \mathrm{M}\}$, where:

π_i denotes the initial probability for state i

$p_{i,j}$ denotes the transition probability between state i and j

μ_i^m denotes the intercept for asset m in state i

$\Sigma_i^{m,n}$ denotes the covariance between asset m and n in state i

Consider the conditional probability function $f(y_t \mid S_t = i, \Theta)$ as the likelihood function for state i conditional on the set of parameters. When the filter probability function $\alpha_{i,t}$ is estimated, one has the necessary information in order to calculate the full log likelihood of the HMM as a function of the set of parameters. The full log likelihood function is given by:

$$ln\ L = \textstyle\sum_{t=1}^{T} ln \sum_{i=1}^{N} f(y_t \mid S_t = i, \Theta)\mathrm{P}(S_t = i) \quad (13)$$

The function (13) is a weighted average of the likelihood function in each state, where the weights are given by the state's probabilities.

One estimates the set of parameters {Θ} by maximising the full log likelihood function over said set of parameters. This must be made under certain conditions, since we are working with probabilities.

Denote the transition probability matrix, i.e. the matrix with element *(i,j)* being the transition probability between state i and j, $p_{i,j}$, by **P**. All elements of this matrix must be non-negative, and all rows must sum up to one. The same goes for the HMM's filter probabilities (Definition 3.1) and the smoothing probabilities (Definition 3.2) (Campigotto, 2009, Perlin, 2015).

6 Practical Implementation Part One: Application to the Stockholm Stock Market

For the application of the model I have chosen two different portfolios, each containing seven stocks selected from the Stockholm stock market in the period 2006-10-02 – 2016-11-01. Said portfolios will be modelled according to the Markov regime switching model given by equation (5) – (10), and the results will be analysed as to how well the selected model fits the data. This will be made by e.g. normalising the residuals after the model has been applied to see how well the time series fits the assumption of normal distributed residuals, along with checking whether any dependence structure remains after normalisation.

One of these portfolios will in the second part on the implementation be evaluated for the last two years of the chosen time period. Using the first year as data base, the expected return and covariance will be estimated from the Markov regime switching model's parameter estimation once every month for the last year of observations. The result will be applied to find the optimal portfolio weights to maximize the Sharpe ratio, and finally compared to an equally weighted portfolio containing the same stocks.

6.1 Definition of Y

The time series of observations constituting the observed part of the HMM gives us a matrix of log returns, denoted **Y**. Each column $\boldsymbol{Y}^m$ of **Y** stands for the log returns of stock m, where element y_t^m denotes the log return of the closing price of stock m in a given portfolio, at time t;

$$y_t^m = ln\left(\frac{c_t^m}{c_{t-1}^m}\right)$$

where c_t^m gives us the closing price of stock m at time t. We work in discrete time, as the data will be daily prices. The observed time period of business days 2006-10-02 – 2016-11-01 of the seven selected stocks makes **Y** a 2535x7 matrix.

Two distinct portfolios each containing seven stocks were chosen from the Stockholm stock market. Portfolio one, containing only stocks with sector root code financials, will be denoted $Port_1$, and portfolio two, with no financial stocks, will be denoted $Port_2$.

Chart 6.1: Asset list for $Port_1$ and $Port_2$

Stock Index	$\boldsymbol{Port_1}$	$\boldsymbol{Port_2}$
$\boldsymbol{Y^1}$	Investor A	AstraZeneca
$\boldsymbol{Y^2}$	Nordea Bank	Volvo A
$\boldsymbol{Y^3}$	SEB A	Hennes & Mauritz B
$\boldsymbol{Y^4}$	Castellum	Alfa Laval
$\boldsymbol{Y^5}$	Fabege	Skanska B
$\boldsymbol{Y^6}$	Handelsbanken A	Axfood
$\boldsymbol{Y^7}$	Swedbank A	PEAB B

Comparing the norms of each portfolio's covariance matrix we get the following results:

$$Port_1: \left\|\begin{matrix} {\sigma^1}^2 & \cdots & \sigma^{1,7} \\ \vdots & \ddots & \vdots \\ \sigma^{7,1} & \cdots & {\sigma^7}^2 \end{matrix}\right\|_2 = 0.002406$$

$$Port_2: \left\| \begin{matrix} \sigma^{1^2} & \cdots & \sigma^{1,7} \\ \vdots & \ddots & \vdots \\ \sigma^{7,1} & \cdots & \sigma^{7^2} \end{matrix} \right\|_2 = 0.001569$$

This demonstrates the reason for selecting one portfolio with financial stocks and one without; the significant difference in the variance and covariance between the stocks of the portfolios. $Port_1$ has a higher norm indicating higher variances and covariances than $Port_2$.

Financial stocks are in general known to be more difficult to model because of the high volatility they usually demonstrate. Therefore, it would be interesting to apply the Markov regime switching model to a portfolio consisting only of financial stocks.

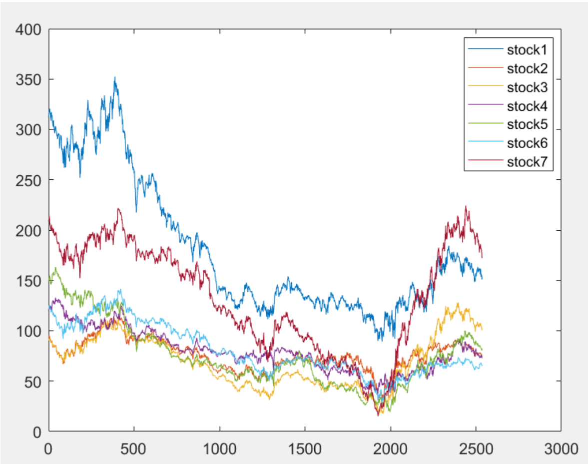

Fig. 6.1 Closing price $Port_1$.

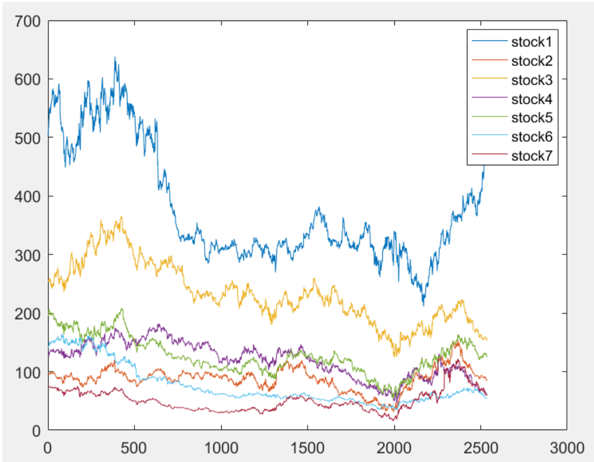

Fig. 6.2 Closing price $Port_2$.

The difference in covariance becomes more apparent when comparing the plots of each portfolio's closing prices (Fig. 6.1, Fig. 6.2). $Port_1$'s larger covariances are evident, as the closing prices of all assets contained in the portfolio appear to move more in sync than those for $Port_2$.

The plots also emphasize the change in the closing prices behaviour. It's clear that the time series follow calmer patterns at some points, while the volatilities increase significantly at other time periods. This argues for the idea of the time series switching between different states, making the Markov regime switching model appealing.

6.2 Parameter Estimation

To estimate the parameter set {Θ} = $\{(\pi_i), (p_{i,j}), (\mu_i^m), (\sigma_i^{n,m})\}$ using the previously defined maximum likelihood formula (13) can be a challenging affair by hand. For this thesis, the MATLAB package: "MS Regress - The MATLAB Package for Markov Regime Switching Models" by Perlin, 2015, was used to estimate the necessary parameters. The package uses the method mentioned in section 5.1, i.e. estimation by using Hamilton's filter to calculate the hidden chain's filter probabilities, before inserting them into the full log likelihood formula (13).

The estimation was made under the assumption of a two- or three-state regime switching model, $S_t \in \Omega$ = {1, …, N} where N = 2 or 3. For simplicity, the remaining part of this section the description of the chosen model will be done for a three-state regime switching model.

The assumption of normal distributed residuals, ε_t, was also made during the estimation.

This presents us with, for each portfolio, a model where the log returns of stock m are drawn from 3 distinct normal distributions, depending on what state the HMM is currently in:

$$Y_t^m = \mu_1^m + \varepsilon_t \quad \text{for state 1} \tag{14}$$

$$Y_t^m = \mu_2^m + \varepsilon_t \quad \text{for state 2} \tag{15}$$

$$Y_t^m = \mu_3^m + \varepsilon_t \quad \text{for state 3} \tag{16}$$

Where;

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_1^{m^2}) \quad \text{for state 1} \tag{17}$$

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_2^{m^2}) \quad \text{for state 2} \tag{18}$$

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_3^{m^2}) \quad \text{for state 3} \tag{19}$$

With seven stocks in each portfolio, we would get seven sets of equation (14) – (19) for each portfolio, with μ_i^m determining the expected log return and $\sigma_i^{m^2}$ the variance for asset m while the HMM is in state i. Since the covariance between assets most likely differs depending on what state the HMM is currently in, one covariance matrix per stock was estimated for each state i;

$$\Sigma_i = \begin{bmatrix} \sigma_i^{1^2} & \cdots & \sigma_i^{1,7} \\ \vdots & \ddots & \vdots \\ \sigma_i^{7,1} & \cdots & \sigma_i^{7^2} \end{bmatrix}, \; i \in \Omega = \{1,2,3\}$$

The expected log returns for state i are presented in a column-vector denoted $\boldsymbol{\mu_i}$, and the residuals in a 2535x7 matrix denoted $\boldsymbol{\varepsilon}$.

The transition probability matrix is in a three-regime model given by:

$$\boldsymbol{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{bmatrix}$$

with $p_{i,j}, \; i,j \epsilon \{1,2,3\}$, defined by equation (11).

As for the remaining parameter set (π_i), i.e. the initial probability of the HMM, it is derived from the smoothing probability mentioned in Definition 3.2. This probability is very similar to the filter probability defined in Definition 3.1, but while the filter probability gives us the probability of the HMM being in a certain state at a time t given all observations $y_{1:t}^m$, the smoothing probability instead determines the probability of the HMM being in a certain state at a previous time. The vector

$$\boldsymbol{\pi}_t = [\pi_{1,t} \;\; \pi_{2,t} \;\; \pi_{3,t}]$$

gives us the corresponding initial probabilities at time t.

6.2.1 Parameter Estimation: Results

Chart 6.2: Two-State Regime Switching Model

$\boldsymbol{Port_1}$	$\boldsymbol{Port_2}$
$\boldsymbol{\mu_1} = \begin{bmatrix} 0.0000 \\ 0.0004 \\ 0.0001 \\ 0.0001 \\ -0.0001 \\ 0.0000 \\ 0.0000 \end{bmatrix} \quad \boldsymbol{\mu_2} = \begin{bmatrix} -0.0008 \\ 0.0001 \\ 0.0001 \\ -0.0002 \\ -0.0004 \\ -0.0004 \\ 0.0000 \end{bmatrix}$	$\boldsymbol{\mu_1} = \begin{bmatrix} 0.0000 \\ 0.0000 \\ 0.0003 \\ 0.0003 \\ 0.0005 \\ 0.0002 \\ 0.0001 \end{bmatrix} \quad \boldsymbol{\mu_2} = \begin{bmatrix} 0.0000 \\ 0.0000 \\ -0.0006 \\ 0.0000 \\ -0.0001 \\ -0.0004 \\ -0.0001 \end{bmatrix}$
$\boldsymbol{\Sigma_1} = \begin{bmatrix} 0.00020 & 0.00018 & 0.00019 & 0.00014 & 0.00014 & 0.00015 & 0.00018 \\ 0.00018 & 0.00027 & 0.00024 & 0.00014 & 0.00015 & 0.00019 & 0.00023 \\ 0.00019 & 0.00024 & 0.00030 & 0.00016 & 0.00017 & 0.00020 & 0.00026 \\ 0.00014 & 0.00014 & 0.00016 & 0.00022 & 0.00019 & 0.00012 & 0.00016 \\ 0.00014 & 0.00015 & 0.00017 & 0.00019 & 0.00027 & 0.00013 & 0.00016 \\ 0.00015 & 0.00019 & 0.00020 & 0.00012 & 0.00013 & 0.00020 & 0.00020 \\ 0.00018 & 0.00023 & 0.00026 & 0.00016 & 0.00016 & 0.00020 & 0.00030 \end{bmatrix}$ $\boldsymbol{\Sigma_2} = \begin{bmatrix} 0.00049 & 0.00041 & 0.00049 & 0.00015 & 0.00019 & 0.00037 & 0.00042 \\ 0.00041 & 0.00107 & 0.00071 & 0.00011 & 0.00012 & 0.00062 & 0.00054 \\ 0.00049 & 0.00071 & 0.00156 & 0.00027 & 0.00041 & 0.00069 & 0.00108 \\ 0.00015 & 0.00011 & 0.00027 & 0.00093 & 0.00079 & 0.00021 & 0.00032 \\ 0.00019 & 0.00012 & 0.00041 & 0.00079 & 0.00136 & 0.00025 & 0.00040 \\ 0.00037 & 0.00062 & 0.00069 & 0.00021 & 0.00025 & 0.00093 & 0.00050 \\ 0.00042 & 0.00054 & 0.00108 & 0.00032 & 0.00040 & 0.00050 & 0.00175 \end{bmatrix}$	$\boldsymbol{\Sigma_1} = \begin{bmatrix} 0.00011 & 0.00005 & 0.00005 & 0.00005 & 0.00005 & 0.00003 & 0.00004 \\ 0.00005 & 0.00027 & 0.00010 & 0.00016 & 0.00013 & 0.00005 & 0.00014 \\ 0.00005 & 0.00010 & 0.00014 & 0.00009 & 0.00008 & 0.00004 & 0.00008 \\ 0.00005 & 0.00016 & 0.00009 & 0.00021 & 0.00012 & 0.00005 & 0.00012 \\ 0.00005 & 0.00013 & 0.00008 & 0.00012 & 0.00015 & 0.00005 & 0.00012 \\ 0.00003 & 0.00005 & 0.00004 & 0.00005 & 0.00005 & 0.00012 & 0.00005 \\ 0.00004 & 0.00014 & 0.00008 & 0.00012 & 0.00012 & 0.00005 & 0.00023 \end{bmatrix}$ $\boldsymbol{\Sigma_2} = \begin{bmatrix} 0.00045 & 0.00006 & 0.00009 & 0.00004 & 0.00010 & 0.00006 & 0.00011 \\ 0.00006 & 0.00103 & 0.00031 & 0.00058 & 0.00052 & 0.00012 & 0.00057 \\ 0.00009 & 0.00031 & 0.00047 & 0.00030 & 0.00028 & 0.00008 & 0.00028 \\ 0.00004 & 0.00058 & 0.00030 & 0.00095 & 0.00046 & 0.00009 & 0.00046 \\ 0.00010 & 0.00052 & 0.00028 & 0.00046 & 0.00073 & 0.00009 & 0.00051 \\ 0.00006 & 0.00012 & 0.00008 & 0.00009 & 0.00009 & 0.00046 & 0.00008 \\ 0.00011 & 0.00057 & 0.00028 & 0.00046 & 0.00051 & 0.00008 & 0.00130 \end{bmatrix}$
$\boldsymbol{P} = \begin{bmatrix} 0.95 & 0.05 \\ 0.18 & 0.82 \end{bmatrix}$	$\boldsymbol{P} = \begin{bmatrix} 0.88 & 0.12 \\ 0.28 & 0.72 \end{bmatrix}$
BIC: -1.049688e+05	BIC: -1.001626e+05

Chart 6.3: Three-State Regime Switching Model

$\boldsymbol{Port_1}$	$\boldsymbol{Port_2}$
$\boldsymbol{\mu_1} = \begin{bmatrix} 0.0004 \\ 0.0000 \\ 0.0000 \\ 0.0004 \\ 0.0002 \\ 0.0003 \\ 0.0000 \end{bmatrix} \boldsymbol{\mu_2} = \begin{bmatrix} -0.0006 \\ -0.0001 \\ -0.0001 \\ -0.0003 \\ -0.0002 \\ -0.0004 \\ 0.0000 \end{bmatrix} \boldsymbol{\mu_3} = \begin{bmatrix} 0.0004 \\ -0.0001 \\ -0.0001 \\ 0.0002 \\ 0.0001 \\ 0.0000 \\ 0.0001 \end{bmatrix}$	$\boldsymbol{\mu_1} = \begin{bmatrix} 0.0000 \\ 0.0001 \\ 0.0003 \\ 0.0005 \\ 0.0002 \\ 0.0003 \\ 0.0003 \end{bmatrix} \boldsymbol{\mu_2} = \begin{bmatrix} 0.0000 \\ -0.0001 \\ -0.0003 \\ -0.0001 \\ -0.0005 \\ 0.0000 \\ 0.0000 \end{bmatrix} \boldsymbol{\mu_3} = \begin{bmatrix} 0.0000 \\ -0.0001 \\ 0.0007 \\ 0.0004 \\ 0.0002 \\ 0.0007 \\ 0.0002 \end{bmatrix}$

$\mathbf{\Sigma_1} = \begin{bmatrix} 0.00024 & 0.00025 & 0.00024 & 0.00014 & 0.00016 & 0.00021 & 0.00021 \\ 0.00025 & 0.00034 & 0.00030 & 0.00015 & 0.00017 & 0.00026 & 0.00026 \\ 0.00024 & 0.00030 & 0.00033 & 0.00015 & 0.00017 & 0.00026 & 0.00027 \\ 0.00014 & 0.00015 & 0.00015 & 0.00016 & 0.00015 & 0.00013 & 0.00014 \\ 0.00016 & 0.00017 & 0.00017 & 0.00015 & 0.00021 & 0.00015 & 0.00015 \\ 0.00021 & 0.00026 & 0.00026 & 0.00013 & 0.00015 & 0.00025 & 0.00023 \\ 0.00021 & 0.00026 & 0.00027 & 0.00014 & 0.00015 & 0.00023 & 0.00027 \end{bmatrix}$ $\mathbf{\Sigma_2} = \begin{bmatrix} 0.00031 & 0.00029 & 0.00034 & 0.00023 & 0.00025 & 0.00025 & 0.00033 \\ 0.00029 & 0.00045 & 0.00044 & 0.00026 & 0.00028 & 0.00033 & 0.00044 \\ 0.00034 & 0.00044 & 0.00061 & 0.00031 & 0.00035 & 0.00038 & 0.00055 \\ 0.00023 & 0.00026 & 0.00031 & 0.00042 & 0.00036 & 0.00023 & 0.00033 \\ 0.00025 & 0.00028 & 0.00035 & 0.00036 & 0.00052 & 0.00025 & 0.00036 \\ 0.00025 & 0.00033 & 0.00038 & 0.00023 & 0.00025 & 0.00035 & 0.00039 \\ 0.00033 & 0.00044 & 0.00055 & 0.00033 & 0.00036 & 0.00039 & 0.00069 \end{bmatrix}$ $\mathbf{\Sigma_3} = \begin{bmatrix} 0.00062 & 0.00046 & 0.00062 & 0.00036 & 0.00043 & 0.00043 & 0.00058 \\ 0.00046 & 0.00157 & 0.00090 & 0.00044 & 0.00048 & 0.00083 & 0.00090 \\ 0.00062 & 0.00090 & 0.00221 & 0.00046 & 0.00054 & 0.00087 & 0.00149 \\ 0.00036 & 0.00044 & 0.00046 & 0.00105 & 0.00077 & 0.00036 & 0.00051 \\ 0.00043 & 0.00048 & 0.00054 & 0.00077 & 0.00140 & 0.00047 & 0.00049 \\ 0.00043 & 0.00083 & 0.00087 & 0.00036 & 0.00047 & 0.00119 & 0.00071 \\ 0.00058 & 0.00090 & 0.00149 & 0.00051 & 0.00049 & 0.00071 & 0.00237 \end{bmatrix}$	$\mathbf{\Sigma_1} = \begin{bmatrix} 0.00009 & 0.00004 & 0.00004 & 0.00004 & 0.00004 & 0.00003 & 0.00004 \\ 0.00004 & 0.00021 & 0.00008 & 0.00013 & 0.00011 & 0.00004 & 0.00011 \\ 0.00004 & 0.00008 & 0.00010 & 0.00007 & 0.00006 & 0.00003 & 0.00006 \\ 0.00004 & 0.00013 & 0.00007 & 0.00017 & 0.00010 & 0.00004 & 0.00010 \\ 0.00004 & 0.00011 & 0.00006 & 0.00010 & 0.00011 & 0.00004 & 0.00009 \\ 0.00003 & 0.00004 & 0.00003 & 0.00004 & 0.00004 & 0.00008 & 0.00004 \\ 0.00004 & 0.00011 & 0.00006 & 0.00010 & 0.00009 & 0.00004 & 0.00017 \end{bmatrix}$ $\mathbf{\Sigma_2} = \begin{bmatrix} 0.00021 & 0.00011 & 0.00010 & 0.00011 & 0.00011 & 0.00004 & 0.00009 \\ 0.00011 & 0.00057 & 0.00024 & 0.00035 & 0.00032 & 0.00012 & 0.00032 \\ 0.00010 & 0.00024 & 0.00030 & 0.00021 & 0.00021 & 0.00007 & 0.00020 \\ 0.00011 & 0.00035 & 0.00021 & 0.00044 & 0.00028 & 0.00011 & 0.00026 \\ 0.00011 & 0.00032 & 0.00021 & 0.00028 & 0.00039 & 0.00010 & 0.00028 \\ 0.00004 & 0.00012 & 0.00007 & 0.00011 & 0.00010 & 0.00029 & 0.00011 \\ 0.00009 & 0.00032 & 0.00020 & 0.00026 & 0.00028 & 0.00011 & 0.00054 \end{bmatrix}$ $\mathbf{\Sigma_3} = \begin{bmatrix} 0.00089 & 0.00012 & 0.00009 & 0.00008 & 0.00021 & 0.00014 & 0.00016 \\ 0.00012 & 0.00165 & 0.00039 & 0.00088 & 0.00081 & 0.00014 & 0.00095 \\ 0.00009 & 0.00039 & 0.00065 & 0.00044 & 0.00036 & 0.00010 & 0.00040 \\ 0.00008 & 0.00088 & 0.00044 & 0.00169 & 0.00074 & 0.00015 & 0.00071 \\ 0.00021 & 0.00081 & 0.00036 & 0.00074 & 0.00117 & 0.00014 & 0.00079 \\ 0.00014 & 0.00014 & 0.00010 & 0.00015 & 0.00014 & 0.00058 & 0.00007 \\ 0.00016 & 0.00095 & 0.00040 & 0.00071 & 0.00079 & 0.00007 & 0.00243 \end{bmatrix}$
$\boldsymbol{P} = \begin{bmatrix} 0.94 & 0.03 & 0.03 \\ 0.06 & 0.90 & 0.04 \\ 0.04 & 0.16 & 0.80 \end{bmatrix}$	$\boldsymbol{P} = \begin{bmatrix} 0.79 & 0.19 & 0.02 \\ 0.20 & 0.74 & 0.06 \\ 0.07 & 0.22 & 0.71 \end{bmatrix}$
BIC: $-1.055782 \cdot 10^5$	BIC: $-1.004746 \cdot 10^5$

6.3 Analysis of Estimation Results

6.3.1 Chart 6.2

Looking at the results of the estimation of the two-state model in Chart 6.2 one can deduce that for both portfolios the expected log return for each stock in state two is always lower or equal to the expected log return in state one. This, along with the significantly higher variance demonstrated by both portfolios in state two hints to the conclusion that state one stands for a "Bull" market environment, with higher returns and lower variances, while state two stands for the opposite; a "Bear" market environment. Note that while the results of this estimation indicate one state being more desirable than the other, this may not always be the case. This will become evident when analysing Chart 6.3.

Comparing the norms of the estimated covariance matrices $\mathbf{\Sigma}_i$ of each portfolio one can find once again that $Port_1$ has higher covariances than $Port_2$, in both states.

Lastly, it is found by observing the transition probability matrix **P** that the regimes in $Port_1$ tend to have a longer duration than the corresponding regimes in $Port_2$. If the HMM is currently in state one, there is a probability of 0.95 that it will remain in this state for $Port_1$, while the same probability is equal to 0.88 for $Port_2$. If the switching between states is too frequent, it is difficult to anticipate the behaviour of the HMM. Therefore, a long regime duration is preferred for most implementations.

To conclude the analysis of Chart 6.2, it would be interesting to relate the portfolios smoothing probability to its conditional standard deviation:

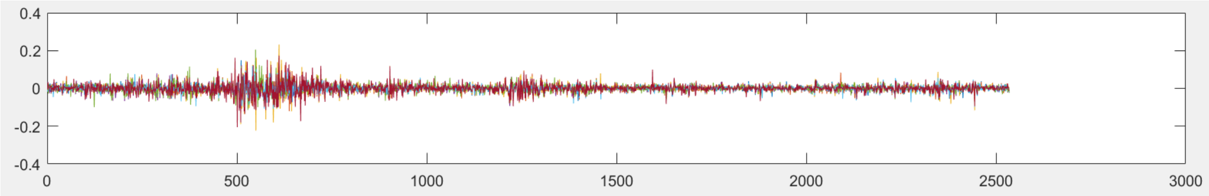

Fig. 6.3 Conditional standard deviation $Port_1$.

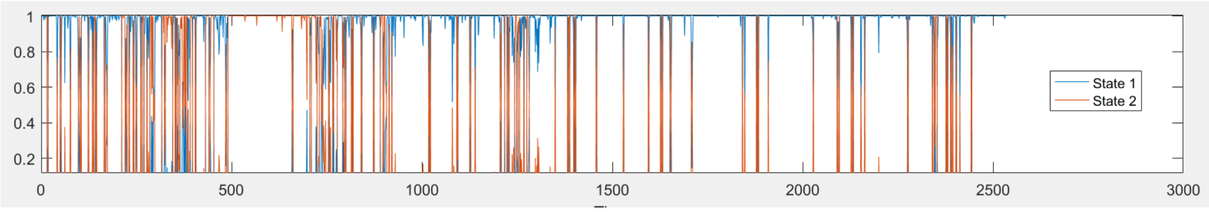

Fig. 6.4 Smoothing probabilities $Port_1$, two states.

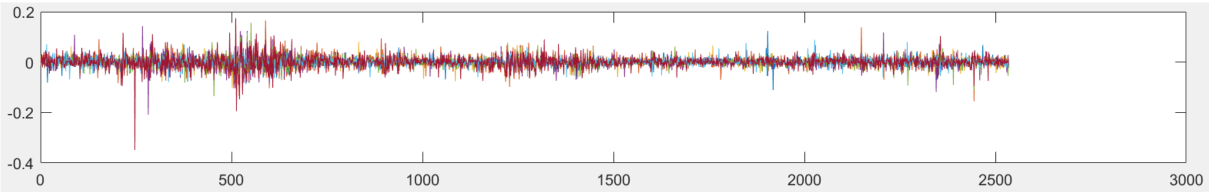

Fig. 6.5 Conditional standard deviation $Port_2$.

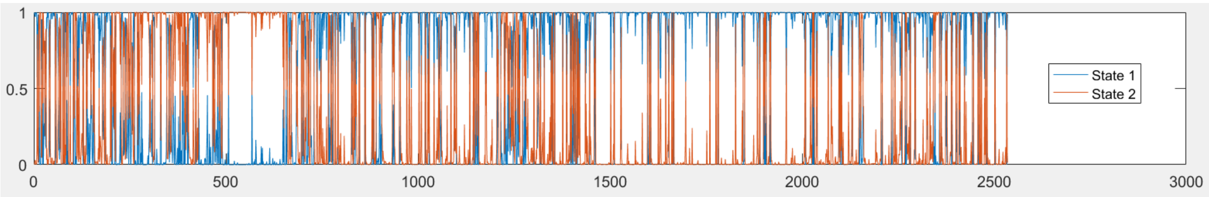

Fig. 6.6 Smoothing probabilities $Port_2$, two states.

Here, one can easily distinguish the difference in regime persistence between the both portfolios. $Port_2$'s smoothing probabilities (Fig. 6.6) indicate an HMM that switches frequently between the two states, while the smoothing probabilities of $Port_1$ (Fig. 6.4) hints to a HMM that stays in each state for a more extensive period of time.

In addition to this, it is also easy to relate the smoothing probabilities to the behaviour of the corresponding conditional standard deviation of each portfolio. Where the higher volatilities of the financial crisis are evident, throughout observation 500 – 700 approximately, the smoothing probabilities of both portfolios demonstrate that state two, the "Bear" regime, is the most probable state of the HMM. For $Port_1$, one can also distinguish the calmer time periods with lower volatilities, consisting of e.g. observation 1900 – 2100 circa, where state one, the "Bull" regime, is the most probable state.

6.3.2 Chart 6.3

When applying the three-state regime switching model to our portfolios the results differ slightly from the two-state scenario.

Firstly, whereas the two-state model had a clear distinction of which state was more preferred to the other, the three-state model displays a less evident separation between states. The norms of the covariance matrices for both portfolios show us that, once again, the regimes are ordered by increasing volatilities. State one is the most desirable state with the lowest variances, while state three has the highest volatilities. This pattern, however, is not followed by the expected log return vectors $\boldsymbol{\mu_1}$ - $\boldsymbol{\mu_3}$. A quick look at $\boldsymbol{\mu_1}$ tells us that state one also has the highest log returns, but for $\boldsymbol{\mu_2}$ and $\boldsymbol{\mu_3}$ the results differ. State two always appear to have an expected log return equal to or lower than for state three. This makes it harder to rank the states in the same way as for the two-state regime switching model, but it doesn't create a problem for this thesis implementation. Generally, one merely has to take risk aversion into account in order to decide whether state two or state three is the less desirable state.

Upon comparison between portfolios the results are similar to the two-state model's: estimated covariances are in general higher for $Port_1$, and the transition probability matrices also indicate a higher regime duration than for $Port_2$.

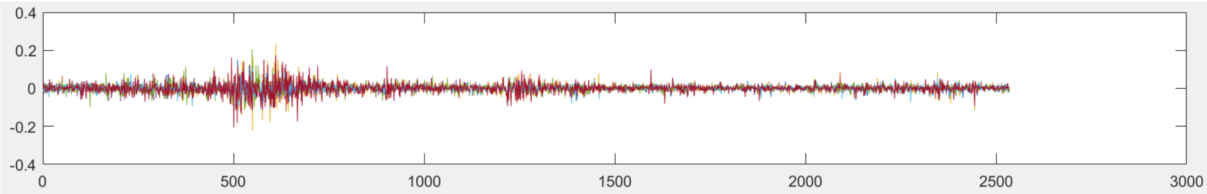

Fig. 6.7 Conditional standard deviation $Port_1$.

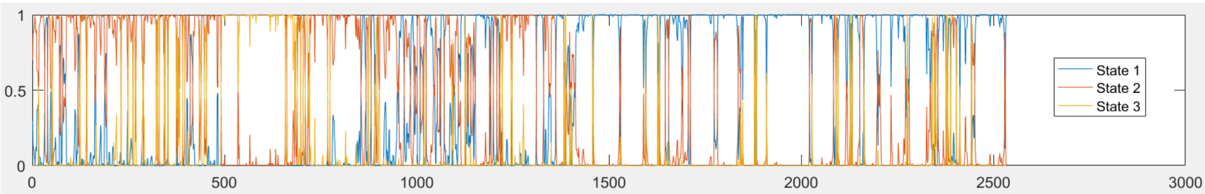

Fig. 6.8 Smoothing probability $Port_1$, three states.

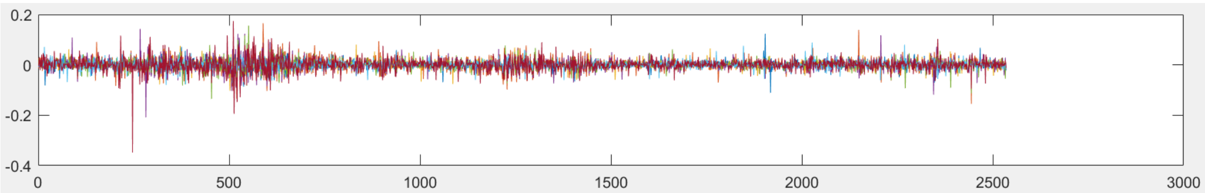

Fig. 6.9 Conditional standard deviation $Port_2$.

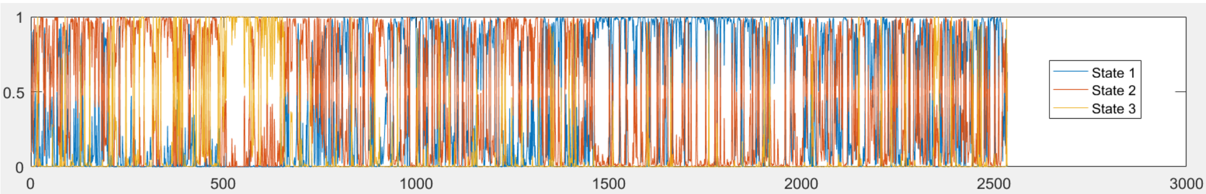

Fig. 6.10 Smoothing probability $Port_2$, three states.

The low regime persistence for Port_2 (Fig. 6.10) is even more evident in the three-state switching model; it is difficult to distinguish what state the HMM is in at all given times in the current scale. Nevertheless, the financial crisis is still noticeable throughout observation 500 – 700 circa, where state three with the highest variances is indubiously the most probable state for both portfolios.

To conclude, the estimated BIC's mentioned in Definition 3.3 indicate that the three-state Markov regime switching model gives the best fit to the observed log return in comparison to the two-state model. This goes for both portfolios; the BIC value for the three-state model is lower than for the two-state model in both cases. Although this reveals which model of the two is preferable by the Bayesian information criterion, it is important to mention that it tells us nothing of how well the model fits the data in general.

A relevant side note is that a four-state regime switching model was also tested for both portfolios, although it brought poorer results; the BIC value had increased in comparison to the three-state model. Further testing established that the four-state model had a worse fit to our data than the other models. In regard to this, only the results of the two- and three-state models will be demonstrated and analysed throughout this thesis.

6.4 Normalisation of Residuals

To see how well the chosen Markov regime switching model fits the selected log return time series in a general sense is obviously of great interest, otherwise one might not be able to trust the results. Apart from comparing BIC values, this will be done by normalising the residuals after the model has been applied to see how well the time series fits the assumption of normal distributed residuals. If the assumption is correct, the following statements should be true:

1) The residuals should not be correlated with another variable.
2) Adjacent residuals should not be correlated with each other, i.e. there should be no autocorrelation between residuals.

If 1) and 2) are not satisfied, there is most likely some explanatory information that hasn't been captured by the model leaking into the residuals, e.g. a missing interaction between terms in the model. If the residuals are autocorrelated there is some predictive information present that is not captured by the predictors.

In order to test if statement 1) and 2) are correct, the estimated residuals of each portfolio will be normalised. Each asset's residuals will be divided by its corresponding standard deviation, conditional on what state the HMM is most probable to be in at the given time. To estimate what state the HMM has the highest probability of being in at time t, the smoothing probability $\pi_{i,t}$ will be taken into account. As mentioned in Definition 3.2, the smoothing probability determines the probability of the HMM being in a given state i at time t, conditionally on the set of observations $\{y_{1:t+\tau}\}, \tau > 0$. In this case, this set will contain all observations included in the estimation, i.e. $\{y_{1:2535}\}$.

If $\pi_{i,t} > 0.5$, the residuals ε_t of all stocks of the given portfolio at time t are assumed to be emitted from state i, and will consequently be divided by the corresponding standard deviation for state i, σ_i. By denoting the 2535x7 matrix of normalised residuals $\boldsymbol{\varepsilon}^{norm}$ we get the expression:

$$\boldsymbol{\varepsilon}^{norm} = \boldsymbol{\varepsilon}\sqrt{\boldsymbol{\Sigma}_i^{-1}} \;, i \in \Omega = \{1, \dots, N\} \tag{20}$$

6.4.1 Normalisation of Residuals: Results and Analysis

Formula (20) was applied to both portfolios, first for the two-state Markov regime switching model, then for the three-state model.

By statement 1) there should be no covariance between the normalised residuals. This means that the covariance matrix of the normalised residuals should be close to the identity matrix of matching size. This was tested by taking the norm of the covariance matrix, denoted by $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}^{norm}}$, while subtracting the identity matrix;

$$\left\| \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}^{norm}} - \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \right\|_2 \tag{21}$$

Chart 6.4: Results from applying formula (21) to $\boldsymbol{\varepsilon}^{norm}$

Number of states	$Port_1$	$Port_2$
2	0.216534	0.125177
3	0.209483	0.106949

Chart 6.4 reveals that covariance between normalised residuals still can be found, indicating that the residuals haven't been sufficiently standardized since a dependence structure still exists. An important observation is, however, that the normalised residuals seem to be more standardized when modelled with a switching model with three states instead of two.

To deduce whether the normalised residuals $\boldsymbol{\varepsilon}^{norm}$ follow a normal distribution they will be displayed in a normal probability plot:

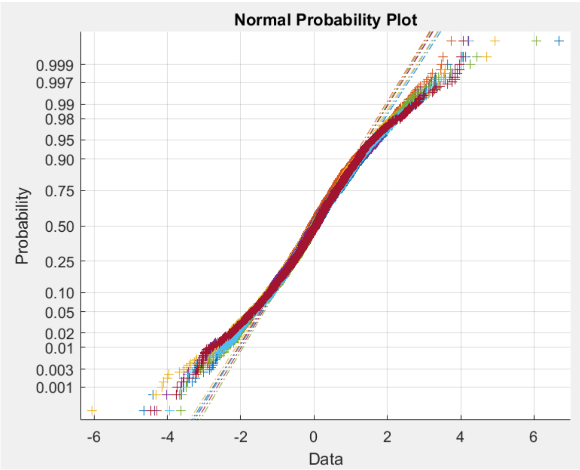

Fig. 6.11 Normal probability plot of normalised residuals $Port_1$, two states.

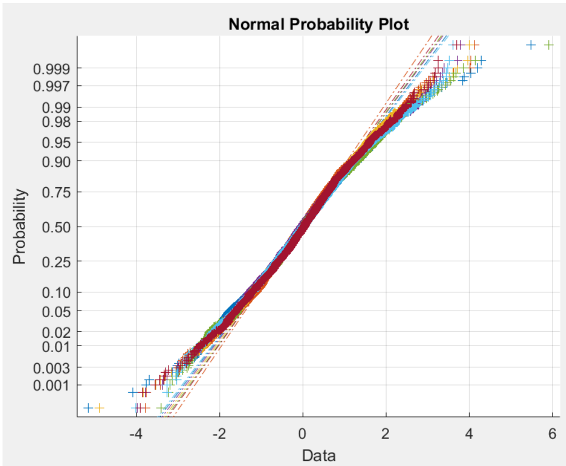

Fig. 6.12 Normal probability plot of normalised residuals $Port_1$, three states.

The plots in Fig. 6.11 and 6.12 demonstrate that $\boldsymbol{\varepsilon}^{norm}$ do not completely follow a normal distribution, which real data is rarely expected to do. When inspecting the ends of the plots they have slopes that are less steep than the fitted line, indicating more outliers. This suggests that the distribution of the normalised residuals has larger tails than a normal distribution. Nonetheless, by closely observing e.g. the axes one can deduce that the normalised residuals of three-state model have fewer extreme outcomes and appear somewhat closer to the fitted line. Thereof, they can be assumed closer to a normal distribution than the normalised residuals of the two-state model.

The next step is to do a scatter plot of $\boldsymbol{\varepsilon}^{norm}$. The aim is for the normalised residuals to look as homogeneous and stationary as possible, meaning there is no dependence structure left between assets. The plots demonstrated in the remaining part of this section will be for $Port_1$, because of the similarity in the results between the portfolios.

Before applying the model one can deduct from Fig. 6.13 that the variance of the log return vector **Y** of $Port_1$ displays a highly fluctuating behaviour, with noticeably higher volatilities e.g. throughout observation 500 – 700.

After applying a two-state Markov regime switching model and normalising the residuals (Fig. 6.14) a significant improvement is evident. The residuals illustrate a relatively homogeneous plot. Yet, upon further inspection of the plot, it would seem that the normalised residuals are slightly further dispersed around observation 500, while they appear more collected e.g. nearby observation 1500 – 2000.

When a three-state model was applied (Fig. 6.15) further improvement was evident. The plot of $Port_1$'s normalised residuals appears to be, if not completely, very close to homogeneous.

Similar results were obtained when the same measures were taken for $Port_2$, only the log return were slightly more homogeneous before any model was applied.

This result shows us that there is no dependence remaining between assets, and consequently between residuals, but only for the same time. Autocorrelation, meaning you can use one residual to predict the next one, is still a possibility. Theoretically, there should be no significant autocorrelation in a financial time series. This would imply that e.g. predicting the future value of a stock would be possible, which it in light of the efficient market hypothesis shouldn't be because of arbitrage.

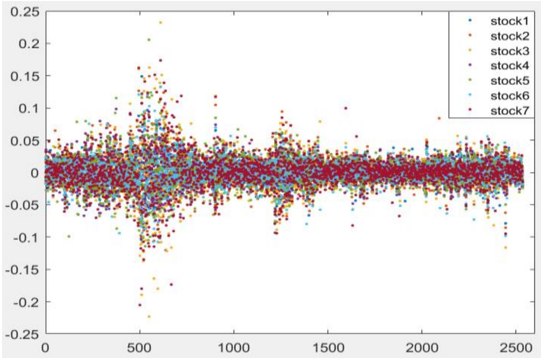

Fig. 6.13 Log return $Port_1$.

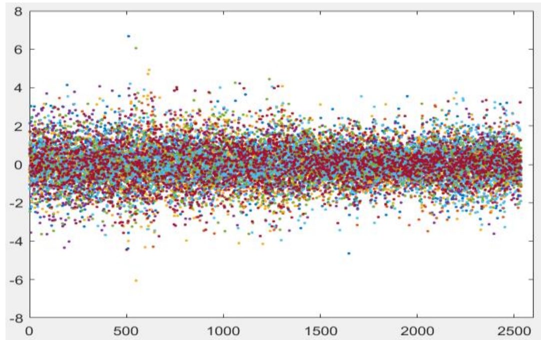

Fig. 6.14 Normalised residuals $Port_1$, two states.

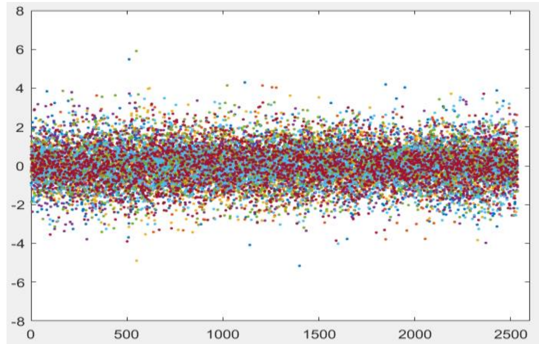

Fig. 6.15 Normalised residuals $Port_1$, three states.

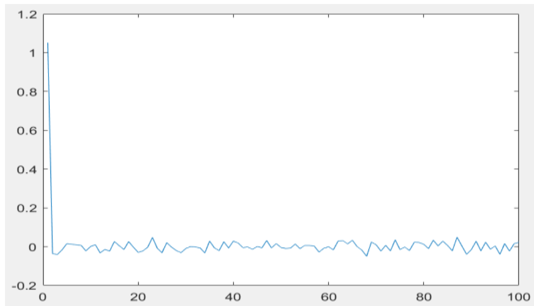

Fig. 6.16 Covariance function of log returns $Port_1$, stock 1.

By observing the covariance function for each stock in the portfolio, one can verify that there seems to be no significant autocovariance, and consequently, no significant autocorrelation. The covariances and correlations in Fig. 6.16-6.23 were estimated with a time interval of 100 business days. As an example, Fig. 6.16 shows us the covariance function of $\boldsymbol{Y}^1$, i.e. stock number one in $Port_1$. It is therefore not surprising that the normalised residuals (Fig. 6.17) too appear to be uncorrelated, regardless of whether they were modelled with the assumption of two or three states.

Matching results were obtained when analysing the remaining stocks in $Port_1$, as well as for all stocks in $Port_2$.

Although time series of log returns often are without any significant autocorrelation, the absolute log returns are not. The autocorrelation function of absolute returns decays slowly with lag, indicating that there is an autocorrelation in the variance of the time series (Granger et al, 2000, Rogers et al, 2011). This is evident when inspecting Fig. 6.13; after a time period of high volatility the returns of the following days are also likely to have high variances.

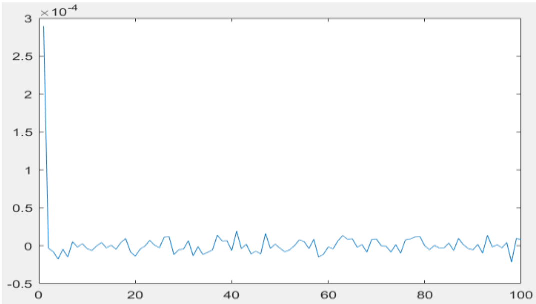

Fig. 6.17 Covariance function of normalised residuals for a two-state model $Port_1$, stock 1.

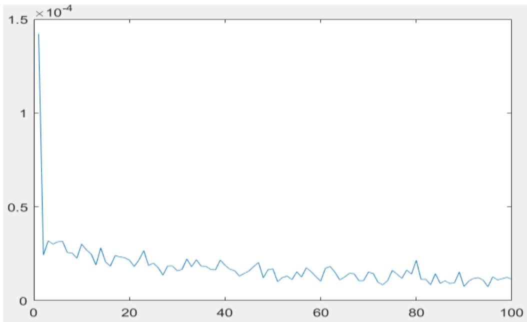

Fig. 6.18 Covariance function of absolute log returns $Port_1$, stock 1.

In order to investigate this, the covariance function of the absolute value of the normalised residuals $\boldsymbol{\varepsilon}^{norm}$ of each stock will be plotted and compared to the covariance function of the corresponding absolute log returns. The trend will also be removed from the plots, enabling us to focus the analysis on the fluctuations in the data.

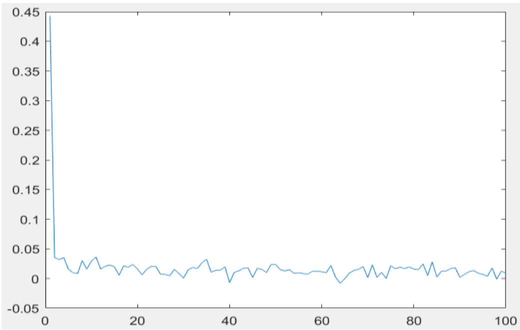

Fig. 6.19 Covariance function of absolute normalised residuals for a two-state model $Port_1$, stock 1.

Fig. 6.18 displays a characteristically decaying covariance function for the absolute log return of an arbitrary stock in $Port_1$, indicating that there is indeed correlation in the variance. After applying either the two- or three-state model and normalising the residuals of either of the portfolios, however, a significant improvement can be distinguished. Only the two-state case is illustrated (Fig. 6.19), since the three-state case presented a nearly identical result.

Now that the autocovariance, and consequently the autocorrelation, of each asset with itself has been checked, it is necessary to investigate whether there is significant autocorrelation between assets. Similar to the result of the log return of a single stock, no significant autocorrelation is detected pairwise between stocks either. Consequently, the autocovariance of pairwise normalised residuals between stocks show no considerable result. Fig. 6.20 illustrates the correlation function between stock six and stock seven in $Port_1$. As predicted, the function fluctuates around zero for all time lags greater than zero. At time lag zero the autocorrelation of each stock with itself reaches the value one, illustrating the autovariance of the stock.

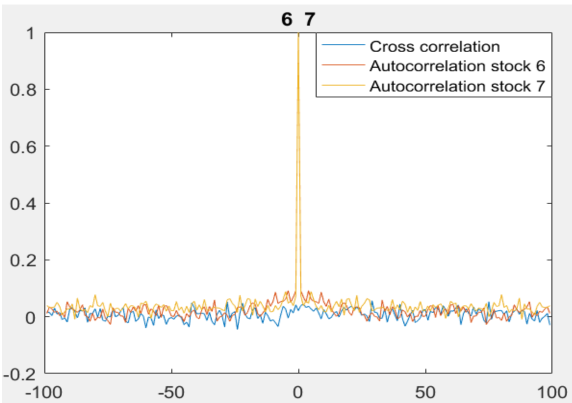

Fig. 6.20 Correlation function between normalised residuals of stock 6 and 7 in $Port_1$, two states.

As for the absolute log return, a dependence that decays with lag is once again found when checking correlation functions pairwise between assets. The correlation function between the absolute log returns of stock six and seven in $Port_1$ (Fig. 6.21) clarifies this. In addition, the cross correlation increases considerably at time lag zero, indicating a high cross correlation between the assets. The correlation functions between all other pairs of stocks in $Port_1$ and $Port_2$ followed the same pattern.

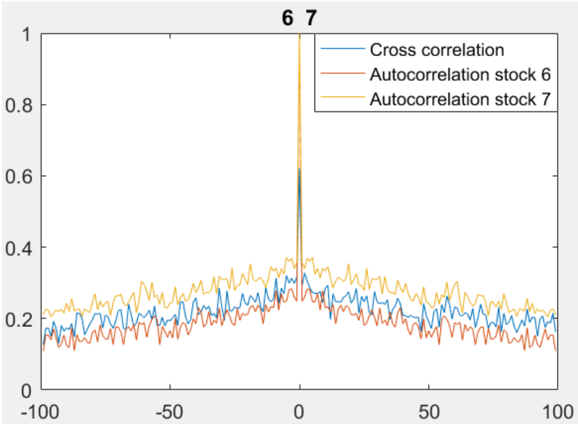

Fig. 6.21 Correlation function between absolute log returns of stock 6 and 7 in $Port_1$.

After applying a two-state regime switching model and normalising the residuals (Fig. 6.22) there is an improvement; the correlation function no longer shows a dependence decaying with lag, and the cross correlation has no increase at time lag zero. Although the improvement is evident, there is still room for further revision. Upon taking a closer look at Fig. 6.22, there is a noticeable gap between the autocorrelations of the stocks. The autocorrelation of stock seven, represented by the yellow line, fluctuates well above zero. This would indicate that there is a dependence that hasn't been captured by the model, but this is not necessarily the case. The irregularity might just indicate a structure breach, e.g. because of the vast difference in variance between distinct time periods. This was a recurring occurrence when analysing pairwise stocks in both portfolios, indicating that the model hasn't successfully captured the entire dependence structure of the data.

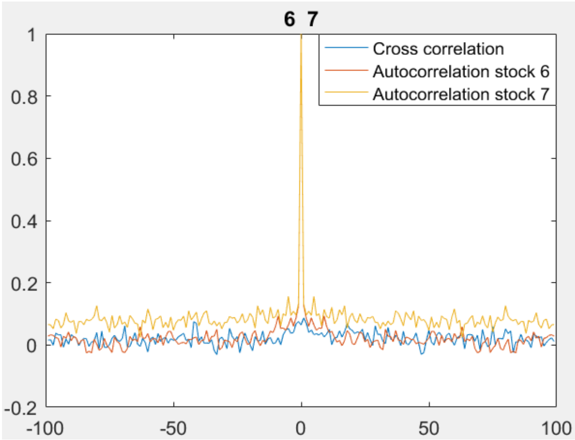

Fig. 6.22 Correlation function between absolute normalised residuals of stock 6 and 7 in $Port_1$, two states.

Applying a three-state regime switching model seemed to remedy this. When applied, all pairwise correlations of absolute normalised residuals showed a more agreeable result (Fig 6.23). The realisations of the autocorrelations were fluctuating around zero, implying that there is no significant dependence structure left after normalisation.

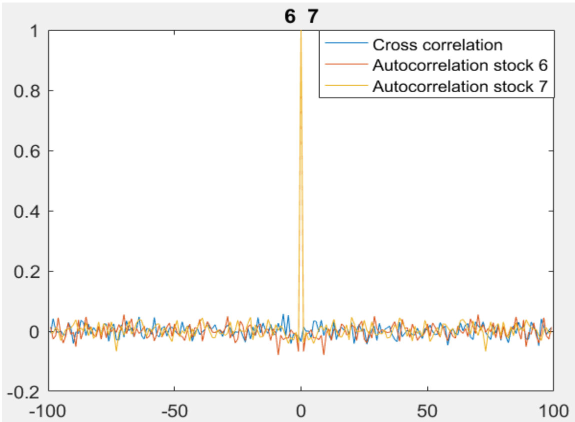

Fig. 6.23 Correlation function between absolute normalised residuals of stock 6 and 7 in $Port_1$, three states.

6.5 Conclusion of Practical Implementation Part One

Implementing a Markov regime switching model to a time series of log returns can be problematic when assuming normal distributed residuals. The data might not fit the assumption, and there might be dependence structures that are not apprehended by the model.

By changing the number of states assumed in the regime switching model it was apparent that one can improve the fit of the model, and it was ultimately found that a three-state model was the best candidate for both portfolios. This was concluded after finding that the three-state model had the lowest BIC value, along with having normalised residuals closest to a normal distribution. The three-state model also attended most effectively to the autocorrelation of the absolute normalised residuals, as it implied that all significant dependence structure was captured by the model.

It should be noted, however, that the normalised residuals did not completely follow a normal distribution. This indicates that the model did not capture the entire behaviour of the time series in a satisfactory fashion.

7 Practical Implementation Part Two: Allocation Optimization

As for the second part of this thesis practical implementation, the model with the best fit to our data will be used to estimate the weights optimizing the Sharpe ratio of the chosen portfolio. The returns of this allocation will then be compared to those of a portfolio of equal weights, containing the same assets.

7.1 Practical Details

The three-state Markov regime switching model was in the first part of the implementation found to fit the log return time series the best. It will be implemented on $Port_1$, since the portfolios residuals appeared just as homogeneous and independent after normalisation as for $Port_2$, in spite of the far higher variations in volatility of the log return before any model was applied.

The column vector of relative weights $\boldsymbol{w}^{max}$ optimizing the Sharpe ratio of $Port_1$ will be estimated each month under the course of one year, using the previous year as base of observations. The last two years of the observations used in part one will be taken into account, i.e. 2014-11-04 – 2016-11-01. This gives us 500 observations of each stock, making the observed log returns **Y** a 500x7 matrix.

The first year, constituted by observation 0-250, will as previously mentioned be used as data base when estimating the necessary parameters. Thereafter, the remaining 250 observations will be partitioned into 12 subseries, one representing each month. Each subseries will contain 20 observations except for the last subseries, which will contain 30, in order to take a full year into account.

For each subseries, the column-vector of expected log returns $\widehat{\boldsymbol{\mu}}$ and the matrix of expected covariances $\widehat{\boldsymbol{\Sigma}}$ will be estimated for the following 20 observations. This estimation will be based on the set of parameters {Θ} obtained when applying a three-state Markov regime switching model. Using these estimates, the relative weights maximising the Sharpe ratio will be estimated, and the absolute weights will then be applied to the actual observed log return. This portfolio will be denoted $Port_{HMM}$. When this is done, a new estimation of the set of necessary parameters for the regime switching model will occur, updating the parameters used to estimate $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ for the successive subseries of observations.

Lastly, the outcome will be compared to a portfolio of equal weights, denoted $Port_{ew}$, containing the same stocks. A hypothetical investment of 1 million Swedish crowns (1 MSEK) will be put into each of the portfolios.

7.2 The Equal Weight Portfolio

Before initializing any estimation, a motivation as to why a portfolio of equal weights was chosen as a fair adversary to the optimal portfolio $Port_{HMM}$ is in place. The equal weight strategy, a.k.a. the naïve strategy, does not involve any optimization or estimation and completely ignores the data, always assigning the relative weight $w^{ew} = \frac{1}{M}$ to each of the M assets.

In spite of the total disregard of the data, it has been found that the equal weight strategy in general has Sharpe ratios that are higher, or statistically indistinguishable, relative to a wide range of strategies. This particular research includes several portfolio strategies, such as the Bayes-Stein shrinkage portfolio, Markowitz mean variance portfolio, etc. The conclusion to this investigation announced that no strategy from the optimal models is consistently better than the equal weight strategy, indicating that the errors in estimating returns and covariances erode all the gains from optimal diversification (DeMiguel et al, 2007).

7.3 Estimation of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$

Given that all necessary parameters: {Θ} = $\{(\pi_i), (p_{i,j}), (\mu_i^m), (\Sigma_i^{m,n})\}$, $\mathrm{i,j} \in \Omega = \{1, 2, \dots, \mathrm{N}\}, \mathrm{m, n} \in \{1, 2, \dots, \mathrm{M}\}$ of a Markov regime switching model have been estimated for a certain sequence of log returns, it is possible to estimate the expected log returns and covariances for a consecutive sequence of not yet observed log returns.

Let $\tilde{Y}_\delta = \sum_{t=1}^{\delta} y_t$ denote the sum of log returns y_t, where δ is the number of steps, or in our case; the number of business days, we want to estimate. One can use the moment generating function of $\tilde{Y}_\delta$ to estimate the requested parameters. By denoting the moment generating function M_δ we get the following formula (Asmussen, 2003):

$$M_\delta(z) = E\left[e^{z \cdot \tilde{Y}_\delta}\right] = \boldsymbol{\pi}_0 \left(\boldsymbol{P} e^{K(z)}\right)^{\delta} \mathbf{1}_N$$

Where;

$\boldsymbol{\pi}_0$ denotes the initial probability of the HMM

$\boldsymbol{P}$ denotes the transition probability matrix

N is the HMM's total number of states

$\mathbf{1}_N$ denotes a size N column-vector of ones

z is a column-vector with elements $z_1, \dots, z_M$, with M denoting the total number of assets

K(z) is a diagonal matrix with elements $z'\boldsymbol{\mu}_1 + z'\boldsymbol{\Sigma}_1 \cdot {}^{z}/_{2}, \dots, z'\boldsymbol{\mu}_N + z'\boldsymbol{\Sigma}_N \cdot {}^{z}/_{2}$, where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ denotes the expected log return vector and covariance matrix of the HMM in state i, respectively.

It is now possible to express the estimated expected log return for the consecutive time period δ of stock m as:

$$\hat{\mu}^m = E\left[e^{\tilde{Y}_\delta^m}\right] - 1 = E\left[e^{\boldsymbol{I}'_m \cdot \tilde{Y}_\delta}\right] - 1 = M_\delta(\boldsymbol{I}_i) - 1$$

Where $\boldsymbol{I}_m$ denotes column m in a MxM identity matrix.

In addition to this, the covariance between asset m and n, i.e. element *(m,n)* in the estimated covariance matrix $\hat{\boldsymbol{\Sigma}}$, can be obtained by the following formula:

$$\hat{\Sigma}^{m,n} = E\left[e^{\tilde{Y}_\delta{}^m + \tilde{Y}_\delta{}^n}\right] - E\left[e^{\tilde{Y}_\delta{}^m}\right] E\left[e^{\tilde{Y}_\delta{}^n}\right] = E\left[e^{(\boldsymbol{I}_m + \boldsymbol{I}_n)' \cdot \tilde{Y}_\delta}\right] - E\left[e^{\boldsymbol{I}_m{}' \cdot \tilde{Y}_\delta}\right] E\left[e^{\boldsymbol{I}_n{}' \cdot \tilde{Y}_\delta}\right]$$
$$= M_\delta(\boldsymbol{I}_m + \boldsymbol{I}_n) - M_\delta(\boldsymbol{I}_m) M_\delta(\boldsymbol{I}_n)$$

7.3.1 Estimation of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$: Results

Chart 7.1: $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ for $Port_1$

Subseries index	$\hat{\boldsymbol{\mu}}$	$\hat{\boldsymbol{\Sigma}}$
1	$\hat{\boldsymbol{\mu}}_1 = \begin{bmatrix} 0.0081 \\ -0.0012 \\ 0.0044 \\ 0.0108 \\ 0.0302 \\ 0.0020 \\ 0.0018 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_1 = \begin{bmatrix} 0.0046 & 0.0035 & 0.0031 & 0.0026 & 0.0029 & 0.0030 & 0.0030 \\ 0.0035 & 0.0054 & 0.0034 & 0.0024 & 0.0028 & 0.0035 & 0.0030 \\ 0.0031 & 0.0034 & 0.0041 & 0.0018 & 0.0021 & 0.0031 & 0.0030 \\ 0.0026 & 0.0024 & 0.0018 & 0.0035 & 0.0030 & 0.0021 & 0.0019 \\ 0.0029 & 0.0028 & 0.0021 & 0.0030 & 0.0043 & 0.0023 & 0.0023 \\ 0.0030 & 0.0035 & 0.0031 & 0.0021 & 0.0023 & 0.0041 & 0.0030 \\ 0.0030 & 0.0030 & 0.0030 & 0.0019 & 0.0023 & 0.0030 & 0.0036 \end{bmatrix}$
2	$\hat{\boldsymbol{\mu}}_2 = \begin{bmatrix} -0.0103 \\ -0.0170 \\ -0.0210 \\ -0.0168 \\ -0.0030 \\ -0.0207 \\ -0.0212 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_2 = \begin{bmatrix} 0.0273 & 0.0259 & 0.0256 & 0.0252 & 0.0258 & 0.0256 & 0.0253 \\ 0.0259 & 0.0277 & 0.0256 & 0.0251 & 0.0255 & 0.0259 & 0.0255 \\ 0.0256 & 0.0256 & 0.0264 & 0.0243 & 0.0248 & 0.0257 & 0.0252 \\ 0.0252 & 0.0251 & 0.0243 & 0.0262 & 0.0261 & 0.0246 & 0.0243 \\ 0.0258 & 0.0255 & 0.0248 & 0.0261 & 0.0275 & 0.0252 & 0.0249 \\ 0.0256 & 0.0259 & 0.0257 & 0.0246 & 0.0252 & 0.0267 & 0.0253 \\ 0.0253 & 0.0255 & 0.0252 & 0.0243 & 0.0249 & 0.0253 & 0.0258 \end{bmatrix}$
3	$\hat{\boldsymbol{\mu}}_3 = \begin{bmatrix} 0.0081 \\ 0.0044 \\ 0.0012 \\ 0.0108 \\ 0.0302 \\ 0.0020 \\ 0.0018 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_3 = \begin{bmatrix} 0.0046 & 0.0035 & 0.0031 & 0.0026 & 0.0029 & 0.0030 & 0.0030 \\ 0.0035 & 0.0054 & 0.0034 & 0.0024 & 0.0028 & 0.0035 & 0.0030 \\ 0.0031 & 0.0034 & 0.0041 & 0.0018 & 0.0021 & 0.0031 & 0.0030 \\ 0.0026 & 0.0024 & 0.0018 & 0.0035 & 0.0030 & 0.0021 & 0.0019 \\ 0.0029 & 0.0028 & 0.0021 & 0.0030 & 0.0043 & 0.0023 & 0.0023 \\ 0.0030 & 0.0035 & 0.0031 & 0.0021 & 0.0023 & 0.0041 & 0.0030 \\ 0.0030 & 0.0030 & 0.0030 & 0.0019 & 0.0023 & 0.0030 & 0.0036 \end{bmatrix}$
4	$\hat{\boldsymbol{\mu}}_4 = \begin{bmatrix} 0.0021 \\ 0.0008 \\ -0.0153 \\ 0.0078 \\ 0.0235 \\ -0.0064 \\ -0.0048 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_4 = \begin{bmatrix} 0.0040 & 0.0033 & 0.0029 & 0.0023 & 0.0025 & 0.0029 & 0.0027 \\ 0.0033 & 0.0052 & 0.0033 & 0.0024 & 0.0025 & 0.0036 & 0.0031 \\ 0.0029 & 0.0033 & 0.0041 & 0.0018 & 0.0020 & 0.0033 & 0.0030 \\ 0.0023 & 0.0024 & 0.0018 & 0.0035 & 0.0030 & 0.0022 & 0.0019 \\ 0.0025 & 0.0025 & 0.0020 & 0.0030 & 0.0041 & 0.0023 & 0.0020 \\ 0.0029 & 0.0036 & 0.0033 & 0.0022 & 0.0023 & 0.0047 & 0.0031 \\ 0.0027 & 0.0031 & 0.0030 & 0.0019 & 0.0020 & 0.0031 & 0.0035 \end{bmatrix}$
5	$\hat{\boldsymbol{\mu}}_5 = \begin{bmatrix} -0.0396 \\ -0.0423 \\ -0.0429 \\ -0.0365 \\ -0.0305 \\ -0.0451 \\ -0.0473 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_5 = \begin{bmatrix} 0.0475 & 0.0467 & 0.0465 & 0.0456 & 0.0463 & 0.0461 & 0.0458 \\ 0.0467 & 0.0487 & 0.0471 & 0.0457 & 0.0462 & 0.0469 & 0.0462 \\ 0.0465 & 0.0471 & 0.0482 & 0.0453 & 0.0458 & 0.0470 & 0.0463 \\ 0.0456 & 0.0457 & 0.0453 & 0.0466 & 0.0466 & 0.0451 & 0.0450 \\ 0.0463 & 0.0462 & 0.0458 & 0.0466 & 0.0484 & 0.0458 & 0.0456 \\ 0.0461 & 0.0469 & 0.0470 & 0.0451 & 0.0458 & 0.0478 & 0.0460 \\ 0.0458 & 0.0462 & 0.0463 & 0.0450 & 0.0456 & 0.0460 & 0.0467 \end{bmatrix}$
6	$\hat{\boldsymbol{\mu}}_6 = \begin{bmatrix} 0.0017 \\ -0.0043 \\ -0.0074 \\ 0.0070 \\ 0.0185 \\ -0.0084 \\ -0.0037 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_6 = \begin{bmatrix} 0.0049 & 0.0036 & 0.0037 & 0.0028 & 0.0031 & 0.0036 & 0.0035 \\ 0.0036 & 0.0055 & 0.0040 & 0.0029 & 0.0030 & 0.0041 & 0.0036 \\ 0.0037 & 0.0040 & 0.0053 & 0.0024 & 0.0027 & 0.0042 & 0.0038 \\ 0.0028 & 0.0029 & 0.0024 & 0.0038 & 0.0033 & 0.0024 & 0.0024 \\ 0.0031 & 0.0030 & 0.0027 & 0.0033 & 0.0046 & 0.0028 & 0.0027 \\ 0.0036 & 0.0041 & 0.0042 & 0.0024 & 0.0028 & 0.0052 & 0.0038 \\ 0.0035 & 0.0036 & 0.0038 & 0.0024 & 0.0027 & 0.0038 & 0.0045 \end{bmatrix}$
7	$\hat{\boldsymbol{\mu}}_7 = \begin{bmatrix} 0.0143 \\ 0.0005 \\ -0.0004 \\ 0.0088 \\ 0.0205 \\ 0.0013 \\ 0.0017 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_7 = \begin{bmatrix} 0.0047 & 0.0039 & 0.0035 & 0.0026 & 0.0029 & 0.0035 & 0.0033 \\ 0.0039 & 0.0059 & 0.0042 & 0.0027 & 0.0028 & 0.0042 & 0.0038 \\ 0.0035 & 0.0042 & 0.0055 & 0.0020 & 0.0023 & 0.0043 & 0.0036 \\ 0.0026 & 0.0027 & 0.0020 & 0.0039 & 0.0034 & 0.0022 & 0.0022 \\ 0.0029 & 0.0028 & 0.0023 & 0.0034 & 0.0049 & 0.0025 & 0.0027 \\ 0.0035 & 0.0042 & 0.0043 & 0.0022 & 0.0025 & 0.0056 & 0.0038 \\ 0.0033 & 0.0038 & 0.0036 & 0.0022 & 0.0027 & 0.0038 & 0.0047 \end{bmatrix}$

8	$\hat{\boldsymbol{\mu}}_8 = \begin{bmatrix} 0.0053 \\ -0.0073 \\ 0.0005 \\ 0.0121 \\ 0.0114 \\ 0.0006 \\ -0.0009 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_8 = \begin{bmatrix} 0.0045 & 0.0038 & 0.0036 & 0.0026 & 0.0026 & 0.0035 & 0.0031 \\ 0.0038 & 0.0058 & 0.0041 & 0.0026 & 0.0026 & 0.0041 & 0.0038 \\ 0.0036 & 0.0041 & 0.0056 & 0.0022 & 0.0021 & 0.0042 & 0.0037 \\ 0.0026 & 0.0026 & 0.0022 & 0.0037 & 0.0030 & 0.0022 & 0.0022 \\ 0.0026 & 0.0026 & 0.0021 & 0.0030 & 0.0042 & 0.0025 & 0.0024 \\ 0.0035 & 0.0041 & 0.0042 & 0.0022 & 0.0025 & 0.0053 & 0.0036 \\ 0.0031 & 0.0038 & 0.0037 & 0.0022 & 0.0024 & 0.0036 & 0.0045 \end{bmatrix}$
9	$\hat{\boldsymbol{\mu}}_9 = \begin{bmatrix} 0.0039 \\ -0.0022 \\ -0.0107 \\ 0.0111 \\ 0.0140 \\ -0.0087 \\ -0.0013 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_9 = \begin{bmatrix} 0.0049 & 0.0039 & 0.0040 & 0.0027 & 0.0031 & 0.0036 & 0.0036 \\ 0.0039 & 0.0065 & 0.0045 & 0.0029 & 0.0030 & 0.0042 & 0.0039 \\ 0.0040 & 0.0045 & 0.0059 & 0.0025 & 0.0028 & 0.0046 & 0.0041 \\ 0.0027 & 0.0029 & 0.0025 & 0.0041 & 0.0035 & 0.0023 & 0.0025 \\ 0.0031 & 0.0030 & 0.0028 & 0.0035 & 0.0049 & 0.0027 & 0.0031 \\ 0.0036 & 0.0042 & 0.0046 & 0.0023 & 0.0027 & 0.0056 & 0.0039 \\ 0.0036 & 0.0039 & 0.0041 & 0.0025 & 0.0031 & 0.0039 & 0.0050 \end{bmatrix}$
10	$\hat{\boldsymbol{\mu}}_{10} = \begin{bmatrix} 0.0136 \\ -0.0007 \\ -0.0094 \\ 0.0094 \\ 0.0175 \\ -0.0015 \\ 0.0024 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_{10} = \begin{bmatrix} 0.0053 & 0.0038 & 0.0039 & 0.0028 & 0.0028 & 0.0036 & 0.0034 \\ 0.0038 & 0.0062 & 0.0043 & 0.0026 & 0.0026 & 0.0042 & 0.0037 \\ 0.0039 & 0.0043 & 0.0055 & 0.0022 & 0.0024 & 0.0043 & 0.0040 \\ 0.0028 & 0.0026 & 0.0022 & 0.0038 & 0.0033 & 0.0021 & 0.0022 \\ 0.0028 & 0.0026 & 0.0024 & 0.0033 & 0.0046 & 0.0023 & 0.0026 \\ 0.0036 & 0.0042 & 0.0043 & 0.0021 & 0.0023 & 0.0054 & 0.0037 \\ 0.0034 & 0.0037 & 0.0040 & 0.0022 & 0.0026 & 0.0037 & 0.0051 \end{bmatrix}$
11	$\hat{\boldsymbol{\mu}}_{11} = \begin{bmatrix} 0.0058 \\ -0.0005 \\ -0.0026 \\ 0.0122 \\ 0.0211 \\ -0.0040 \\ 0.0023 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_{11} = \begin{bmatrix} 0.0047 & 0.0036 & 0.0036 & 0.0026 & 0.0028 & 0.0035 & 0.0032 \\ 0.0036 & 0.0059 & 0.0042 & 0.0027 & 0.0026 & 0.0042 & 0.0038 \\ 0.0036 & 0.0042 & 0.0057 & 0.0023 & 0.0024 & 0.0045 & 0.0038 \\ 0.0026 & 0.0027 & 0.0023 & 0.0039 & 0.0034 & 0.0022 & 0.0024 \\ 0.0028 & 0.0026 & 0.0024 & 0.0034 & 0.0046 & 0.0025 & 0.0027 \\ 0.0035 & 0.0042 & 0.0045 & 0.0022 & 0.0025 & 0.0055 & 0.0037 \\ 0.0032 & 0.0038 & 0.0038 & 0.0024 & 0.0027 & 0.0037 & 0.0047 \end{bmatrix}$
12	$\hat{\boldsymbol{\mu}}_{12} = \begin{bmatrix} 0.0136 \\ -0.0005 \\ 0.0036 \\ 0.0208 \\ 0.0278 \\ 0.0040 \\ 0.0057 \end{bmatrix}$	$\hat{\boldsymbol{\Sigma}}_{12} = \begin{bmatrix} 0.0068 & 0.0053 & 0.0053 & 0.0036 & 0.0041 & 0.0051 & 0.0048 \\ 0.0053 & 0.0085 & 0.0061 & 0.0035 & 0.0037 & 0.0062 & 0.0054 \\ 0.0053 & 0.0061 & 0.0080 & 0.0031 & 0.0036 & 0.0064 & 0.0057 \\ 0.0036 & 0.0035 & 0.0031 & 0.0054 & 0.0044 & 0.0031 & 0.0030 \\ 0.0041 & 0.0037 & 0.0036 & 0.0044 & 0.0065 & 0.0036 & 0.0037 \\ 0.0051 & 0.0062 & 0.0064 & 0.0031 & 0.0036 & 0.0081 & 0.0057 \\ 0.0048 & 0.0054 & 0.0057 & 0.0030 & 0.0037 & 0.0057 & 0.0068 \end{bmatrix}$

7.4 Estimation of $\boldsymbol{w}^{max}$

The estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ announced in Chart 7.1 are subsequently used to, for each subseries, find the relative portfolio weights $\boldsymbol{w}^{max}$ that maximize the Sharpe ratio S_p of the selected portfolio:

$$S_p = \frac{\mu_p - \mu_f}{\sigma_p} \qquad (22)$$

Where;

μ_p denotes the expected return of portfolio p

μ_f denotes the risk-free rate of return

σ_p denotes the standard deviation of portfolio p

For simplicity, this implementation will have the risk-free rate of return set to zero, meaning that equation (22) will simply be composed of the expected return divided by the standard deviation of the portfolio. In addition, no short selling is allowed, making all weights positive. When conducting pure investment the positive weight strategy is often preferred, e.g. because of the impossibility of losing more capital than the invested amount.

Denote the column-vector of relative portfolio weights $\boldsymbol{w}$. The aim is to find the weights that maximize the Sharpe ratio. Expressing (22) in matrix form, with a zero-risk-free rate return, one wants to maximize the following formula with respect to $\boldsymbol{w}$:

$$max_{\boldsymbol{w}}\, S_p = max_{\boldsymbol{w}}\, \frac{\boldsymbol{w}'\widehat{\boldsymbol{\mu}}}{\sqrt{\boldsymbol{w}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{w}}}$$

All weights need to sum up to one and be positive, so the following conditions are imposed on the elements of $\boldsymbol{w}$:

$$\sum_{i=1}^{M} w_i = 1$$

$$w_i \geq 0, \;\; i \,\epsilon\, \{1, \dots, M\}$$

With M denoting the total number of assets.

7.4.1 Estimation of $\boldsymbol{w}^{max}$: Results

Chart 7.2: $\boldsymbol{w}^{max}$ (transposed) for $Port_1$

Subseries Index	$\boldsymbol{w}^{max}$
1	$\boldsymbol{w}_1^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
2	$\boldsymbol{w}_2^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
3	$\boldsymbol{w}_3^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
4	$\boldsymbol{w}_4^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
5	$\boldsymbol{w}_5^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
6	$\boldsymbol{w}_6^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
7	$\boldsymbol{w}_7^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
8	$\boldsymbol{w}_8^{max'} = [0.000\ 0.000\ 0.000\ 0.760\ 0.240\ 0.000\ 0.000]$
9	$\boldsymbol{w}_9^{max'} = [0.000\ 0.000\ 0.000\ 0.235\ 0.765\ 0.000\ 0.000]$
10	$\boldsymbol{w}_{10}^{max'} = [0.195\ 0.000\ 0.000\ 0.000\ 0.805\ 0.000\ 0.000]$
11	$\boldsymbol{w}_{11}^{max'} = [0.000\ 0.000\ 0.000\ 0.000\ 1.000\ 0.000\ 0.000]$
12	$\boldsymbol{w}_{12}^{max'} = [0.000\ 0.000\ 0.000\ 0.191\ 0.809\ 0.000\ 0.000]$

7.5 Implementation of Estimates

The parameters $\widehat{\boldsymbol{\mu}}$, $\widehat{\boldsymbol{\Sigma}}$ and $\boldsymbol{w}^{max}$ estimated in this section will set the foundation for the allocation of our optimal portfolio $Port_{HMM}$.

Before implementing the optimal weights, the absolute weights must be calculated for each subseries. These are the weights that will be implemented on the observed return. Denoting the optimal absolute returns of asset m at time t h_t^m, this is done by applying the following formula:

$$h_t^m = \frac{P_t \cdot \boldsymbol{I}_t \boldsymbol{w}^{max\prime}}{c_t^m} \qquad (23)$$

Where:

c_t^m denotes the closing price of stock m at time t

P_t denotes the compound return of the investment at time t

Equivalent calculations are done for the equal weight portfolio, $Port_{ew}$, only the relative weights are at all times equal to $w^{ew} = \frac{1}{M}$ for each of the M assets.

The optimal weights given by (23) will be calculated for each subseries and the allocations will be updated accordingly. A hypothetical amount of 1 MSEK will be invested into the first allocation, and the entire return will then be reinvested one month apart in every subsequent allocation. All along, the same investment will be put into $Port_{ew}$. The log returns and compound returns of both portfolios will be announced and compared throughout the entire year.

7.5.1 Implementation of Estimates: Results

Chart 7.3: $Port_{HMM}$ vs. $\boldsymbol{Port_{ew}}$

Subseries Index	*Log Return* $\boldsymbol{Port_{HMM}}$	*Compound Return SEK* $\boldsymbol{Port_{HMM}}$	*Log Return* $\boldsymbol{Port_{ew}}$	*Compound return SEK* $\boldsymbol{Port_{ew}}$
1	0.0098	1 009 772	0.0045	1 004 493
2	-0.1258	882 736	-0.0058	998 665
3	0.1365	1 003 257	0.0429	1 041 490
4	0.0227	1 026 059	0.0142	1 056 276
5	0.1238	1 153 094	0.0559	1 115 323
6	0.0452	1 205 212	0.0061	1 122 147
7	0.0108	1 218 241	-0.0151	1 105 229
8	-0.0272	1 185 144	0.0156	1 122 432
9	0.0166	1 204 784	-0.0168	1 103 585
10	-0.0318	1 166 529	-0.0471	1 051 604
11	-0.0572	1 099 779	-0.0152	1 035 576
12	-0.0575	1 036 544	-0.0455	988 422

7.6 Implementation of Estimates: Analysis

Chart 7.3 presents the obtained results when allocating $Port_{HMM}$ according to our estimated absolute weights. To get an overview of the progress of our investment, Fig. 7.1 shows us the development of the compound return of both portfolios.

According to Chart 7.3, $Port_{HMM}$ ended up with a yearly return of approximately 3.65 %, while $Port_{ew}$ held a negative yearly return of circa -1.16 %. One could thereof reach the conclusion that the portfolio modelled with a three-state regime switching model outperformed the equal weight portfolio.

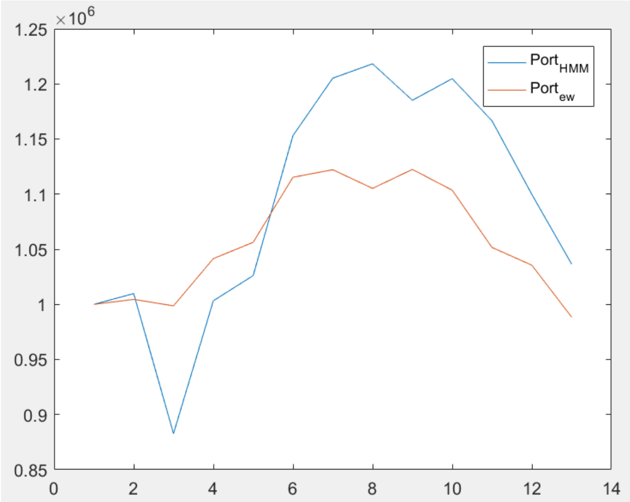

Fig. 7.1 Compound return (MSEK) of $Port_{HMM}$ and $Port_{ew}$.

Upon taking a more sceptical look at the results, however, it is found that the log returns of $Port_{HMM}$ show significantly larger fluctuations, indicating higher volatilities. This, of course, presents considerably higher risks, and a risk avert investor would likely refrain from such an investment.

A closer look at Fig. 7.1 declares that although the compound return of $Port_{HMM}$ at most observed times is decidedly higher than that of $Port_{ew}$, a large dive can be noted at time two, i.e. at the second observed month. When consulting the chart, a significant negative return of -12.58 % can be deduced at the given time. The dive is all but recovered in the third subseries, when a return of 13.65 % is presented. Nevertheless, this substantiates the previous statement of high volatilities.

It would have been interesting to include a comparison of the portfolios Sharpe ratios, but this proved difficult to implement as the data points of the evaluation are too few.

An important side note is the extreme allocations demonstrated by the Sharpe ratio-maximising relative weights $\boldsymbol{w}^{max}$ presented in Chart 7.2. Going by the chart, all weight is to be put into one single asset at most times. A diversifying strategy, i.e. a portfolio constructed of different assets, is often found favourable. The reason for this is that a more diversified portfolio, on average, pose a lower risk than any individual investment found within the portfolio.

In brief, it seems that according to this evaluation the results are inconclusive. In spite of the on average larger returns of $Port_{HMM}$, the considerably smaller fluctuations of the returns of the equal weight portfolio makes it more attractive to a risk avert investor. It is uncertain whether the implementation of the Markov regime switching model led to a more favourable portfolio allocation. To increase the credibility in the results, this could perhaps be remedied by e.g. a larger sample.

8 Conclusions and Discussion

In the first part of the implementation of the Markov regime switching model to the Stockholm stock market it was demonstrated that a three-state model was preferred over a two-state model. This was indicated by the fact that the three-state model had a lower BIC value, along with the overall better fit to the time series of observed log returns. The three-state model also attended most efficiently to the autocorrelation of the absolute normalised residuals, implying that all significant dependence structure was captured by the model. In addition to this, the normalised residuals of the three-state model proved slightly closer to a normal distribution.

However, by observing a normal probability plot of the normalised residuals, one can also conclude that said residuals were not entirely normal distributed. Real data often demonstrate this type of complication, since observed residuals rarely act completely in accordance with any distribution. This presents a complication, since the Markov regime switching model assumes normal distributed residuals when estimating the parameters. Consequently, the chosen Markov regime switching model might not have an as satisfactory fit to the data as assumed in this thesis, leading to a decrease in credibility in the results.

Part two of the implementation showed us that regarding yearly return, the portfolio modelled by a three-state Markov regime switching model outperformed the equal weights-portfolio. As to which portfolio was preferred over the other, however, the results were uncertain. $Port_{HMM}$ displayed significantly larger volatilities exposing the investor to larger risks.

Worth to mention is the size of the implemented subseries; since a year consisted of 250 observations, each month, save for the last, was represented by 20 observations for simplicity. Looking at a real almanac, this is of course not the case.

Another issue with the implementation was the extreme allocations that were estimated for $Port_{HMM}$. Putting one's entire investment into one asset increases the exposure to the particular assets risk, which is why a more diversified strategy is to be preferred.

Despite of apparent complications, we cannot escape the fact that the implementation of the Markov regime switching model led to a considerable improvement in the autocorrelation of the absolute residuals. These results, along with the increased homogeneity of said residuals indicates that most of the significant dependence structure has been captured, in particular by the three-state model. The poor results found in the second part of the implementation may have been due to a too small sample size.

8.1 Difficulties and Further Research

Due to the assumption of normal distributed residuals not being entirely fulfilled, all significant information may not have been caught by the model. A similar study, with e.g. an assumption of generalized error distributed residuals, might yield more trustworthy results and would therefore be interesting to read.

In acknowledgement of the inconclusiveness of the second implementation, a larger sample could with advantage be taken into regard. Because of the amount of computation time involved, this was not possible to implement for this thesis.

9 References

Asmussen, S. (2003). “Applied Probability and Queues”, Stochastic Modelling and Applied Probability, Volume 51, p. 311.

Campigotto, P-J. (2009). “Markov-switching models in Foreign Exchange - how to benefit from trends in the market?” Master’s thesis. Mathematical Statistics, Lund University.

Cont, R. (2001). ”Empirical properties of asset returns: stylized facts and statistical Issues”, Quantitative Finance, 1, p. 223-236.

DeMiguel, V., Garlappi, L., Uppal, R. (2007). “Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?” Oxford University Press.

Dempster, A., Laird, N., and Dubin, D. (1977). “Maximum likelihood from incomplete data via the EM algorithm”, Journal of the Royal Statistical Society, 39, p. 1-38.

Granger, C. W. J., Spear, S., Ding, Z. (2000). “Stylized facts on the temporal and distributional properties of absolute returns: An update” Statistics and Finance: An Interface.

Hamilton, J.D. (1989). “A new approach to the economic analysis of nonstationary time series and the business cycle” Econometrica, 57, p. 357–384

Hamilton, J.D. (1994). “Time Series Analysis” Princeton University Press.

Hamilton, J.D. (2005). “Regime-Switching Models” The New Palgrave Dictionary of Economics, Second Edition.

Lindgren, G. (1978). “Markov Regime Models for Mixed Distributions and Switching Regressions” Scandinavian Journal of Statistics, 5, p. 81-91.

Perlin, M. (2015). “MS Regress - The MATLAB Package for Markov Regime Switching Models” Available at SSRN: http://ssrn.com/abstract=1714016 or http://dx.doi.org/10.2139/ssrn.1714016

Rogers, L.C.G, Zhang, L. (2011). “Understanding Asset Returns” Mathematics and Financial Economics.

Rydén, T., Cappé, O., Moulines, E., (2005). “Inference in Hidden Markov Models” Springer Series in Statistics, First Edition.

Populärvetenskaplig Sammanfattning

Finansiella tidsseriers oförutsägbara beteende har länge varit ett bekymmer för ekonometriker. Det är en utmaning att hitta lämpliga modeller som tar hänsyn till strukturskiftningarna som uppstår i tidsseriernas uppträdande. En Markov regim switching modell är ett populärt redskap, mycket tack vare det sätt den behandlar förändringar i en tidsseries beteende.

Anta att en tidsserie kan beskrivas med hjälp av olika regimer, eller tillstånd, vilka tilldelas allteftersom tidsseriens beteende förändras. Bytesmekanismen som bestämmer övergången från en regim till en annan styrs av en gömd Markov modell. En gömd Markov modell är en statistisk modell som används för att modellera processer, där man tänker sig att ett system kan beskrivas som en mängd tillstånd vilka man rör sig mellan enligt en Markovprocess. Modellen i denna avhandling är baserad på en blandning av normalfördelningar, mellan vilka tidserien rör sig vid regimbyte. Enligt egenskaper hämtade från denna modell så beror tidsseriens aktuella värde endast på det absolut senaste värdet. Det innebär att en struktur i en tidsseries beteende kan råda under en slumpmässig tidsperiod, innan den blir ersatt av en annan struktur när regimbytet sker. På detta sätt kan en Markov regim switching modell fånga mycket komplexa dynamiska mönster.

Modellen applicerades på två portföljer, bestående av vardera sju aktier från Stockholmsbörsen. Efter åtskilliga examinationer var det tydligt att modellens lämplighet, främst utifrån hur väl den passade de utvalda finansiella tidsserierna, kunde förbättras genom att ändra antalet regimer i modellen. Bland annat kontrollerades hur väl de standardiserade residualerna följde en normalfördelning, samt hurvida någon beroendestruktur återfanns hos residualerna efter normalisering. Det konstaterades följaktligen att en Markov regim switching modell med tre regimer var mest fördelaktig. Anpassningen till portföljernas tidsserier var dock inte optimal.

Modellen med tre regimer implementerades således på en av portföljerna, som sedan allokerades i syfte att maximera Sharpe kvoten, som mäter den riskjusterade avkastningen. Detta ledde till extrema portföljvikter, med allt kapital investerat i en enda aktie vid de flesta tidpunkter, vilket anses ofördelaktigt ur riskspridningssynpunkt. Vid jämförelse med likaviktad portfölj, innehållande samma aktier, blev resultaten dessvärre inte tillfredställande. Trots en högre årlig avkastning påvisade den modellerade portföljen betydligt högre volatilitet. Det var alltså inte möjligt att avgöra huruvida appliceringen av en Markov regim switching modell ledde till en mer gynnsam allokering i en underökning av föreliggande omfattning.

Trots uppenbara komplikationer medförde appliceringen av modellen en betydande förbättring rörande residualernas beroendestruktur. Autokorrelationen i den absoluta avkastningen minskade avsevärt, samtidigt som de standardiserade residualerna fick ett mer homogent utseende. Detta leder till slutsatsen att det mesta av tidsseriernas beroendestruktur fångats upp, i synnerhet av Markov regim switching modellen med tre regimer. Implementeringens svaga resultat kan möjligtvis förklaras av att den utfördes över ett för litet urval.