A regime switching model for smart beta investing using hidden Markov models

Elizabeth Fons

University of Manchester, AllianceBernstein

2019 Quant Summit Europe
A novel dynamic asset allocation system using Feature Saliency Hidden Markov models for smart beta investing

Elizabeth Fons¹,³ Paula Dawson¹ Jeffrey Yau¹,² Xiao-jun Zeng³ John Keane³

¹AllianceBernstein, London, UK.
²University of California, Berkeley
³University of Manchester

Abstract

The financial crisis of 2008 generated interest in more transparent, rules-based strategies for portfolio construction, with Smart beta strategies emerging as a novel approach involving asset allocation. While this approach is not new, its popularity among investors and financial practitioners has increased in recent years. However, the majority of research on regime-switching models in finance is that it considers either

Keywords: Hidden Markov model, Dynamic asset allocation, Portfolio optimization, Feature Selection, Smart Beta

1. Introduction

Smart beta is a relatively new term that has become ubiquitous in asset management, particularly for institutional investors. While the concept has been around since the 1980s, there have been more recent developments in the last few years, and understanding whether any additional risk is required for a portfolio is an important consideration.

By selecting stocks based on their factor exposures, active managers can build portfolios with higher than market returns, risk, and underperformance, with more risk focused portfolios showing some improvement. In a first step we build a dynamic asset allocation model using the Feature Saliency Hidden Markov (FSHMM) algorithm that simultaneously performs feature selection and regime identification.

While smart beta strategies have been promising, they are not without their challenges. One of the key challenges is how to incorporate regime information from the HMM, to improve regime identification. We follow a systematic approach in regime switching frameworks for asset allocation.

The primary advantage of regime switching models is that they can dynamically adjust to changing market conditions, allowing for improved portfolio performance. In this study, we evaluate the performance of the FSHMM algorithm on a set of MSCI indices and find that it performs better than a single regime allocation model.

The results show that the FSHMM algorithm can improve portfolio performance by up to 60% in excess of the market annually. In addition, we propose a novel smart beta allocation system based on the FSHMM algorithm that incorporates regime information into the asset allocation process. The results show that the FSHMM algorithm can improve return, and understanding whether any additional risk is added to the portfolio is an important consideration.

Finally, the common factor in the majority of the research and innovation programme under the Marie Sklodowska-Curie grant agreement No 675044

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 675044
Overview

1 Motivation

2 Part 1
   ▶ Hidden Markov Models
   ▶ Data
   ▶ DAA system and Implementation details
   ▶ Main results

3 Part 2
   ▶ Feature selection and HMMs
   ▶ Implementation details
   ▶ Main results

4 Conclusions and future work
Motivation

Factor investing is becoming more relevant each year, especially factor rotation and multi-factor strategies. Hidden Markov models have been used extensively on many financial problems but not a lot of research on their application to factor investing.
Motivation

Factor investing is becoming more relevant each year, especially factor rotation and multi-factor strategies. Hidden Markov models have been used extensively on many financial problems but not a lot of research on their application to factor investing.

Two main contributions:
Motivation

Factor investing is becoming more relevant each year, especially factor rotation and multi-factor strategies. Hidden Markov models have been used extensively on many financial problems but not a lot of research on their application to factor investing.

Two main contributions:

1. Use HMMs to identify market regimes and to build the portfolios. We did a systematic study with multiple factors and types of portfolios.
Motivation

Factor investing is becoming more relevant each year, especially factor rotation and multi-factor strategies. Hidden Markov models have been used extensively on many financial problems but not a lot of research on their application to factor investing.

Two main contributions:

1. Use HMMs to identify market regimes and to build the portfolios. We did a systematic study with multiple factors and types of portfolios.

2. Use an unsupervised feature selection algorithm to select optimal features for the HMM. We show that feature selection improves regime identification.
Hidden Markov Models

Unsupervised clustering method, very useful for sequential data, since it’s able to handle temporal correlations.
Hidden Markov Models

Unsupervised clustering method, very useful for sequential data, since it’s able to handle temporal correlations.

\(X\): sequence of latent states (market regime in our case) that can’t be observed directly and are modeled as a Markov chain.

\(A\) and \(\pi_0\): transition matrix and initial state distribution.

\(\mu\) and \(\sigma\): mean and std deviation of Gaussian distribution.
Hidden Markov Models

Unsupervised clustering method, very useful for sequential data, since it’s able to handle temporal correlations.

\( \mathbf{X} \): sequence of latent states (market regime in our case) that can’t be observed directly and are modeled as a Markov chain.

\( \mathbf{Y} \): sequence of observed data (factor index returns in our case).
Hidden Markov Models

Unsupervised clustering method, very useful for sequential data, since it’s able to handle temporal correlations.

**X**: sequence of latent states (market regime in our case) that can’t be observed directly and are modeled as a Markov chain.

**Y**: sequence of observed data (factor index returns in our case).

Parameters:

- **A** and **π₀**: transition matrix and initial state distribution.
- **μ** and **σ**: mean and std deviation of Gaussian distribution.
Hidden Markov Models

Unsupervised clustering method, very useful for sequential data, since it’s able to handle temporal correlations.

**X**: sequence of latent states (market regime in our case) that can’t be observed directly and are modeled as a Markov chain.

**Y**: sequence of observed data (factor index returns in our case).

Parameters:

- **A** and **π₀**: transition matrix and initial state distribution.
- **μ** and **σ**: mean and std deviation of Gaussian distribution.

\[
\begin{align*}
\{y_{11}, y_{12}, y_{12}, \ldots, y_{1n}\} \\
\{y_{21}, y_{22}, y_{22}, \ldots, y_{2n}\} \\
\vdots \\
\{y_{L1}, y_{L2}, y_{L2}, \ldots, y_{Ln}\}
\end{align*}
\]

\[
\begin{align*}
\text{EM} & \quad A \\
& \quad \pi \\
& \quad \sigma \\
& \quad \mu
\end{align*}
\]
Hidden Markov Models

Unsupervised clustering method, very useful for sequential data, since it’s able to handle temporal correlations.

**X**: sequence of latent states (market regime in our case) that can’t be observed directly and are modeled as a Markov chain.

**Y**: sequence of observed data (factor index returns in our case).

Parameters:

- **A** and **π₀**: transition matrix and initial state distribution.
- **μ** and **σ**: mean and std deviation of Gaussian distribution.

![Diagram showing a Hidden Markov Model](image)
Hidden Markov Model: Example

**Model**

- Input: 3 factor returns
- Nr of states: 3
- Covar: full
Hidden Markov Model: Example

Model
Input: 3 factor returns
Nr of states: 3
Covar: full
Hidden Markov Model: Example

**Model**

Input: 3 factor returns
Nr of states: 3
Covar: full

![Diagram showing the model structure with states S0, S1, and S2 connected with transition probabilities 0.91, 0.95, and 0.99.]

**Table showing factor returns for Value, Quality, and Momentum with states and time periods 1995 to 2015.**
Hidden Markov Model: Example

Model
Input: 3 factor returns
Nr of states: 3
Covar: full

\[
S_0 = \begin{pmatrix} 0.0015 \\ -0.0001 \\ -0.0023 \end{pmatrix} \\
S_1 = \begin{pmatrix} 0.0001 \\ 0.0001 \end{pmatrix} \\
S_2 = \begin{pmatrix} -0.0003 \\ 0.0001 \\ 0.0004 \end{pmatrix}
\]

Value
Quality
Momentum
Data

**Daily factor data from S&P 500**
The universe is ranked, a portfolio is constructed with the top 20% of stocks and short positions in the bottom 20% of stocks.

**MSCI USA enhanced indices**
Start from 6 MSCI indices (Value, Momentum, Size, Volatility, High Yield, Quality) and subtract equally weighted index.
Data

**Daily factor data from S&P500**
The universe is ranked, a portfolio is constructed with the top 20% of stocks and short positions in the bottom 20% of stocks.

**MSCI USA enhanced indices**
Start from 6 MSCI indices (Value, Momentum, Size, Volatility, High Yield, Quality) and subtract equally weighted index.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Date</th>
<th>Nr of features</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor data</td>
<td>Jan-1988 to Feb-2016</td>
<td>25</td>
<td>Daily</td>
</tr>
<tr>
<td>MSCI Enhanced</td>
<td>Jan-1999 to Feb-2016</td>
<td>6</td>
<td>Daily</td>
</tr>
</tbody>
</table>
Dynamic Allocation System

Objective: build a DAA system incorporating information from a HMM and test it systematically.

- State is determined daily, by adding one observation per day (yes, this is very noisy!).
- Given a change of state, use information from model (mean and variance) to recalculate the weights of the portfolio.
- After one month, add the observations to the stack of previous ones with an expanding window and retrain the model.
Calibration and test of DAA system

1. Training
   - Full data
     - Randomly select of 5 factors (one per family)
     - Train HMM

2. Validation
   - Use different delays d to determine change of state
   - Portfolio 4
     - d-4
     - d-44

3. Testing
   - Use DAA system
   - Select optimal d. Keep retraining monthly
   - Portfolio performance
   - Select optimal delay d
Calibration and test of DAA system

1. **Training**
   - Full data
   - Randomly select of 5 factors (one per family)
   - Train HMM

2. **Validation**
   - Use different delays $d$ to determine change of state
     - $d=4$
     - $d=44$
     - Portfolio$_4$
     - Portfolio$_{44}$
   - Portfolio performance
   - Select optimal delay $d$

Testing

Use DAA system: optimal $d$. Keep retraining monthly.
Calibration and test of DAA system

1. Training
   - Full data → Randomly select of 5 factors (one per family) → Train HMM

2. Validation
   - Use different delays $d$ to determine change of state
     - $d=4$ → Portfolio$_4$
     - $d=44$ → Portfolio$_{44}$
   - Portfolio performance → Select optimal delay $d$

3. Testing
   - Use DAA system optimal $d$. Keep retraining monthly.
Implementation details

- We used the python library *hmmlearn* (derived from scikit-learn).
  
  https://hmmlearn.readthedocs.io
Implementation details

- We used the python library *hmmlearn* (derived from scikit-learn).
  
  https://hmmlearn.readthedocs.io

- After model selection, we chose a 2-state model.
Implementation details

- We used the python library *hmmlearn* (derived from scikit-learn).
  [https://hmmlearn.readthedocs.io](https://hmmlearn.readthedocs.io)
- After model selection, we chose a 2-state model.
- Training set corresponds to 15 years (estimating model parameters).
Implementation details

- We used the python library *hmmlearn* (derived from scikit-learn).
  https://hmmlearn.readthedocs.io
- After model selection, we chose a 2-state model.
- Training set corresponds to 15 years (estimating model parameters).
- Validation set corresponds to 9 years (this is used to determine the number of states and when we flag a change of state).
Implementation details

- We used the python library *hmmlearn* (derived from scikit-learn).
  
  https://hmmlearn.readthedocs.io

- After model selection, we chose a 2-state model.

- Training set corresponds to 15 years (estimating model parameters).

- Validation set corresponds to 9 years (this is used to determine the number of states and when we flag a change of state).

- Test set is 4 years.
Implementation details

- We used the python library `hmmlearn` (derived from scikit-learn).
  https://hmmlearn.readthedocs.io
- After model selection, we chose a 2-state model.
- Training set corresponds to 15 years (estimating model parameters).
- Validation set corresponds to 9 years (this is used to determine the number of states and when we flag a change of state).
- Test set is 4 years.
- The model is retrained monthly after the initial training, to account for changes in the parameters (to relax stationarity).
Methodology and baselines

- Risk parity ($\sigma$)
- Max diversification ($\sigma$)
- Min Variance ($\sigma$)

Three groups of portfolios:

- HMM: Six portfolios built using information from HMMs.
- Benchmark 1: Six portfolios rebalanced monthly, single regime.
- Benchmark 2: Equally weighted portfolio rebalanced monthly.

Total = (6 + 6 + 1) * 1260 portfolios.
Methodology and baselines

- Risk parity ($\sigma$)
- Max diversification ($\sigma$)
- Min Variance ($\sigma$)
- Sharpe ($\mu$ and $\sigma$)
Methodology and baselines

- Risk parity ($\sigma$)
- Max diversification ($\sigma$)
- Min Variance ($\sigma$)
- Sharpe ($\mu$ and $\sigma$)
- Max return ($\mu$) Given a vector of means, maximizes return given a constrain that no asset can have a weight greater than 80%.
- Dynamic ($\mu$) If all mean returns are positive, it weights the assets proportional to their mean, else, it equally weights them.
Methodology and baselines

- Risk parity ($\sigma$)
- Max diversification ($\sigma$)
- Min Variance ($\sigma$)
- Sharpe ($\mu$ and $\sigma$)
- Max return ($\mu$) Given a vector of means, maximizes return given a constrain that no asset can have a weight greater than 80%.
- Dynamic ($\mu$) If all mean returns are positive, it weights the assets proportional to their mean, else, it equally weights them.

Three groups of portfolios:

**HMM** Six portfolios built using information from HMMs.

**Benchmark 1** Six portfolios rebalanced monthly, single regime.

**Benchmark 2** Equally weighted portfolio rebalanced monthly.

Total = (6 + 6 + 1) * 1260 portfolios.
Main results part 1

- **Blue:** HMM portfolios
- **Orange:** no-regime portfolios
- **Green:** EQ portfolio.

![Graph showing Sortino ratio for different portfolios](image-url)
## Main results part 1

<table>
<thead>
<tr>
<th>Model</th>
<th>Ann ret</th>
<th>Ann vol</th>
<th>IR</th>
<th>Skw</th>
<th>kurt</th>
<th>D. risk</th>
<th>Sortino</th>
<th>DD</th>
<th>DD days</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ</td>
<td>0.77</td>
<td>2.88</td>
<td>0.26</td>
<td>-0.14</td>
<td>0.81</td>
<td>2.05</td>
<td>0.37</td>
<td>379</td>
<td>318</td>
</tr>
<tr>
<td>Dyn HMM</td>
<td>1.67</td>
<td>4.73</td>
<td>0.34</td>
<td>-0.19</td>
<td>1.35</td>
<td>3.37</td>
<td>0.48</td>
<td>32</td>
<td>291</td>
</tr>
<tr>
<td>Dyn Bench</td>
<td>-0.60</td>
<td>3.98</td>
<td>-0.14</td>
<td>-0.40</td>
<td>1.68</td>
<td>2.96</td>
<td>-0.19</td>
<td>1136</td>
<td>682</td>
</tr>
<tr>
<td>Sharpe HMM</td>
<td>2.31</td>
<td>4.66</td>
<td>0.53</td>
<td>-0.19</td>
<td>1.16</td>
<td>3.29</td>
<td>0.75</td>
<td>429</td>
<td>253</td>
</tr>
<tr>
<td>Sharpe Bench</td>
<td>-3.14</td>
<td>4.89</td>
<td>-0.64</td>
<td>-0.79</td>
<td>4.49</td>
<td>3.80</td>
<td>-0.82</td>
<td>1375</td>
<td>873</td>
</tr>
<tr>
<td>MR HMM</td>
<td>3.19</td>
<td>7.03</td>
<td>0.46</td>
<td>-0.19</td>
<td>1.34</td>
<td>4.98</td>
<td>0.65</td>
<td>35</td>
<td>264</td>
</tr>
<tr>
<td>MR Bench</td>
<td>-5.03</td>
<td>7.20</td>
<td>-0.69</td>
<td>-0.78</td>
<td>3.71</td>
<td>5.63</td>
<td>-0.88</td>
<td>4000</td>
<td>1001</td>
</tr>
<tr>
<td>MV HMM</td>
<td>0.61</td>
<td>2.41</td>
<td>0.24</td>
<td>-0.14</td>
<td>0.96</td>
<td>1.72</td>
<td>0.35</td>
<td>662</td>
<td>309</td>
</tr>
<tr>
<td>MV Bench</td>
<td>-0.12</td>
<td>2.24</td>
<td>-0.07</td>
<td>-0.11</td>
<td>0.83</td>
<td>1.61</td>
<td>-0.09</td>
<td>520</td>
<td>511</td>
</tr>
<tr>
<td>MD HMM</td>
<td>0.69</td>
<td>2.54</td>
<td>0.26</td>
<td>-0.14</td>
<td>1.01</td>
<td>1.80</td>
<td>0.37</td>
<td>340</td>
<td>306</td>
</tr>
<tr>
<td>MD Bench</td>
<td>0.01</td>
<td>2.39</td>
<td>-0.02</td>
<td>-0.12</td>
<td>0.84</td>
<td>1.71</td>
<td>-0.02</td>
<td>454</td>
<td>447</td>
</tr>
<tr>
<td>RP HMM</td>
<td>0.63</td>
<td>2.58</td>
<td>0.24</td>
<td>-0.13</td>
<td>1.04</td>
<td>1.84</td>
<td>0.34</td>
<td>212</td>
<td>302</td>
</tr>
<tr>
<td>RP Bench</td>
<td>0.20</td>
<td>2.40</td>
<td>0.07</td>
<td>-0.13</td>
<td>1.04</td>
<td>1.72</td>
<td>0.10</td>
<td>475</td>
<td>416</td>
</tr>
</tbody>
</table>
## Main results part 1

<table>
<thead>
<tr>
<th></th>
<th>Ann ret</th>
<th>Ann vol</th>
<th>IR</th>
<th>Skw</th>
<th>kurt</th>
<th>D. risk</th>
<th>Sortino</th>
<th>DD</th>
<th>DD days</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ</td>
<td>0.77</td>
<td>2.88</td>
<td>0.26</td>
<td>-0.14</td>
<td>0.81</td>
<td>2.05</td>
<td>0.37</td>
<td>379</td>
<td>318</td>
</tr>
<tr>
<td>Dyn HMM</td>
<td>1.67</td>
<td>4.73</td>
<td>0.34</td>
<td>-0.19</td>
<td>1.35</td>
<td>3.37</td>
<td>0.48</td>
<td>32</td>
<td>291</td>
</tr>
<tr>
<td>Dyn Bench</td>
<td>-0.60</td>
<td>3.98</td>
<td>-0.14</td>
<td>-0.40</td>
<td>1.68</td>
<td>2.96</td>
<td>-0.19</td>
<td>1136</td>
<td>682</td>
</tr>
<tr>
<td>Sharpe HMM</td>
<td>2.31</td>
<td>4.66</td>
<td>0.53</td>
<td>-0.19</td>
<td>1.16</td>
<td>3.29</td>
<td>0.75</td>
<td>429</td>
<td>253</td>
</tr>
<tr>
<td>Sharpe Bench</td>
<td>-3.14</td>
<td>4.89</td>
<td>-0.64</td>
<td>-0.79</td>
<td>4.49</td>
<td>3.80</td>
<td>-0.82</td>
<td>1375</td>
<td>873</td>
</tr>
<tr>
<td>MR HMM</td>
<td>3.19</td>
<td>7.03</td>
<td>0.46</td>
<td>-0.19</td>
<td>1.34</td>
<td>4.98</td>
<td>0.65</td>
<td>35</td>
<td>264</td>
</tr>
<tr>
<td>MR Bench</td>
<td>-5.03</td>
<td>7.20</td>
<td>-0.69</td>
<td>-0.78</td>
<td>3.71</td>
<td>5.63</td>
<td>-0.88</td>
<td>&gt;4000</td>
<td>1001</td>
</tr>
<tr>
<td>MV HMM</td>
<td>0.61</td>
<td>2.41</td>
<td>0.24</td>
<td>-0.14</td>
<td>0.96</td>
<td>1.72</td>
<td>0.35</td>
<td>662</td>
<td>309</td>
</tr>
<tr>
<td>MV Bench</td>
<td>-0.12</td>
<td>2.24</td>
<td>-0.07</td>
<td>-0.11</td>
<td>0.83</td>
<td>1.61</td>
<td>-0.09</td>
<td>520</td>
<td>511</td>
</tr>
<tr>
<td>MD HMM</td>
<td>0.69</td>
<td>2.54</td>
<td>0.26</td>
<td>-0.14</td>
<td>1.01</td>
<td>1.80</td>
<td>0.37</td>
<td>340</td>
<td>306</td>
</tr>
<tr>
<td>MD Bench</td>
<td>0.01</td>
<td>2.39</td>
<td>-0.02</td>
<td>-0.12</td>
<td>0.84</td>
<td>1.71</td>
<td>-0.02</td>
<td>454</td>
<td>447</td>
</tr>
<tr>
<td>RP HMM</td>
<td>0.63</td>
<td>2.58</td>
<td>0.24</td>
<td>-0.13</td>
<td>1.04</td>
<td>1.84</td>
<td>0.34</td>
<td>212</td>
<td>302</td>
</tr>
<tr>
<td>RP Bench</td>
<td>0.20</td>
<td>2.40</td>
<td>0.07</td>
<td>-0.13</td>
<td>1.04</td>
<td>1.72</td>
<td>0.10</td>
<td>475</td>
<td>416</td>
</tr>
</tbody>
</table>
## Main results part 1

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Ann ret</th>
<th>Ann vol</th>
<th>IR</th>
<th>Skw</th>
<th>kurt</th>
<th>D. risk</th>
<th>Sortino</th>
<th>DD</th>
<th>DD days</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ</td>
<td>0.77</td>
<td>2.88</td>
<td>0.26</td>
<td>-0.14</td>
<td>0.81</td>
<td>2.05</td>
<td>0.37</td>
<td>379</td>
<td>318</td>
</tr>
<tr>
<td>Dyn HMM</td>
<td>1.67</td>
<td>4.73</td>
<td>0.34</td>
<td>-0.19</td>
<td>1.35</td>
<td>3.37</td>
<td>0.48</td>
<td>32</td>
<td>291</td>
</tr>
<tr>
<td>Dyn Bench</td>
<td>-0.60</td>
<td>3.98</td>
<td>-0.14</td>
<td>-0.40</td>
<td>1.68</td>
<td>2.96</td>
<td>-0.19</td>
<td>1136</td>
<td>682</td>
</tr>
<tr>
<td>Sharpe HMM</td>
<td>2.31</td>
<td>4.66</td>
<td>0.53</td>
<td>-0.19</td>
<td>1.16</td>
<td>3.29</td>
<td>0.75</td>
<td>429</td>
<td>253</td>
</tr>
<tr>
<td>Sharpe Bench</td>
<td>-3.14</td>
<td>4.89</td>
<td>-0.64</td>
<td>-0.79</td>
<td>4.49</td>
<td>3.80</td>
<td>-0.82</td>
<td>1375</td>
<td>873</td>
</tr>
<tr>
<td>MR HMM</td>
<td>3.19</td>
<td>7.03</td>
<td>0.46</td>
<td>-0.19</td>
<td>1.34</td>
<td>4.98</td>
<td>0.65</td>
<td>35</td>
<td>264</td>
</tr>
<tr>
<td>MR Bench</td>
<td>-5.03</td>
<td>7.20</td>
<td>-0.69</td>
<td>-0.78</td>
<td>3.71</td>
<td>5.63</td>
<td>-0.88</td>
<td>&gt;4000</td>
<td>1001</td>
</tr>
<tr>
<td>MV HMM</td>
<td>0.61</td>
<td>2.41</td>
<td>0.24</td>
<td>-0.14</td>
<td>0.96</td>
<td>1.72</td>
<td>0.35</td>
<td>662</td>
<td>309</td>
</tr>
<tr>
<td>MV Bench</td>
<td>-0.12</td>
<td>2.24</td>
<td>-0.07</td>
<td>-0.11</td>
<td>0.83</td>
<td>1.61</td>
<td>-0.09</td>
<td>520</td>
<td>511</td>
</tr>
<tr>
<td>MD HMM</td>
<td>0.69</td>
<td>2.54</td>
<td>0.26</td>
<td>-0.14</td>
<td>1.01</td>
<td>1.80</td>
<td>0.37</td>
<td>340</td>
<td>306</td>
</tr>
<tr>
<td>MD Bench</td>
<td>0.01</td>
<td>2.39</td>
<td>-0.02</td>
<td>-0.12</td>
<td>0.84</td>
<td>1.71</td>
<td>-0.02</td>
<td>454</td>
<td>447</td>
</tr>
<tr>
<td>RP HMM</td>
<td>0.63</td>
<td>2.58</td>
<td>0.24</td>
<td>-0.13</td>
<td>1.04</td>
<td>1.84</td>
<td>0.34</td>
<td>212</td>
<td>302</td>
</tr>
<tr>
<td>RP Bench</td>
<td>0.20</td>
<td>2.40</td>
<td>0.07</td>
<td>-0.13</td>
<td>1.04</td>
<td>1.72</td>
<td>0.10</td>
<td>475</td>
<td>416</td>
</tr>
</tbody>
</table>
Part 2: Feature selection
Part 2: Feature selection

- Traditionally, features to train HMMs are selected in advance (either expert knowledge or data availability). However, these features don’t necessarily contribute to regime identification, which is our goal.
- In ML it is customary to use feature selection to improve model performance.
- However unsupervised feature selection research for HMMs is limited.

We implemented an embedded feature selection algorithm based on:

Feature Saliency HMM

The idea is to divide features into two groups,
- state dependent (good features), modelled by Gaussians that are state dependent with parameters $\mu_i$ and $\sigma_i$
- state independent (irrelevant features), modelled by gaussians with parameters $\tau$ and $\epsilon$.

$\rho$ (feature saliency) is the probability that the feature is relevant.
Feature Saliency HMM

The idea is to divide features into two groups,
- state dependent (good features), modelled by Gaussians that are state dependent with parameters $\mu_i$ and $\sigma_i$
- state independent (irrelevant features), modelled by gaussians with parameters $\tau$ and $\epsilon$.

$\rho$ (feature saliency) is the probability that the feature is relevant.
Relevant Features

- $k$ is a hyper-parameter that can be used as cost of the feature. The higher its value, the less features will be selected.
**Relevant Features**

- *k* is a hyper-parameter that can be used as cost of the feature. The higher its value, the less features will be selected.

- We can see that the selected parameters are stable, and after a value of *k* above 850, the number of features doesn’t change much.
Relevant Features

- $k$ is a hyper-parameter that can be used as cost of the feature. The higher its value, the less features will be selected.

- We can see that the selected parameters are stable, and after a value of $k$ above 850, the number of features doesn’t change much.

- Selected factors are: Book Value Yield, 1 Yr Fwd Earnings Yield, Sales Yield, 6 Month Price Momentum, 12 Month Price Momentum, EPSCV, Beta.
To test the DAA system incorporating the FS algorithm, we built two cases:

- Trained one HMM with all 25 features.
- Trained one HMM with the relevant subset of features.

FSHMM tends to be more sensible to the distress state - it spends 24% of the time in this state versus 20% of the model trained with full set of features.
DAA system with FSHMM

To test the DAA system incorporating the FS algorithm, we built two cases:

- Trained one HMM with all 25 features.
- Trained one HMM with the relevant subset of features.

FSHMM tends to be more sensible to the distress state - it spends 24% of the time in this state versus 20% of the model trained with full set of features.

Same DAA system as before, with daily evaluation of the state and monthly retraining.

- For the allocation, we used MSCI factor indices, so we had to estimate the mean and covariance for each regime.
- We built Dynamic, Max return and Sharpe portfolios only.
Main results part 2

Cumulated return
Main results part 2

Cumulated return

MR FSHMM
MR HMM
Dyn FSHMM
Sharpe FSHMM
Dyn HMM
Sharpe HMM
## Main results part 2

<table>
<thead>
<tr>
<th>Model</th>
<th>Ann ret</th>
<th>Ann vol</th>
<th>IR</th>
<th>Skw</th>
<th>kurt</th>
<th>D. risk</th>
<th>Sortino</th>
<th>DD</th>
<th>DD days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharpe FSHMM</td>
<td>0.061</td>
<td>0.50</td>
<td>0.12</td>
<td>-0.71</td>
<td>2.85</td>
<td>0.37</td>
<td>0.16</td>
<td>-94</td>
<td>387</td>
</tr>
<tr>
<td>Sharpe HMM</td>
<td>-0.11</td>
<td>0.65</td>
<td>-0.16</td>
<td>-0.70</td>
<td>3.84</td>
<td>0.49</td>
<td>-0.22</td>
<td>-164</td>
<td>522</td>
</tr>
<tr>
<td>Sharpe Bench</td>
<td>-1.62</td>
<td>0.92</td>
<td>-1.76</td>
<td>-2.75</td>
<td>15.0</td>
<td>0.82</td>
<td>-1.98</td>
<td>19825</td>
<td>1452</td>
</tr>
<tr>
<td>Dyn FSHMM</td>
<td>0.39</td>
<td>0.65</td>
<td>0.61</td>
<td>-0.41</td>
<td>0.84</td>
<td>0.47</td>
<td>0.84</td>
<td>-52</td>
<td>141</td>
</tr>
<tr>
<td>Dyn HMM</td>
<td>-0.02</td>
<td>0.60</td>
<td>-0.03</td>
<td>-1.12</td>
<td>9.03</td>
<td>0.45</td>
<td>-0.04</td>
<td>-175</td>
<td>566</td>
</tr>
<tr>
<td>Dyn Bench</td>
<td>-1.10</td>
<td>1.03</td>
<td>-1.07</td>
<td>-2.76</td>
<td>16.2</td>
<td>0.88</td>
<td>-1.24</td>
<td>-1508</td>
<td>1123</td>
</tr>
<tr>
<td>MR FSHMM</td>
<td>2.02</td>
<td>3.20</td>
<td>0.63</td>
<td>-0.39</td>
<td>1.83</td>
<td>2.30</td>
<td>0.88</td>
<td>-82</td>
<td>62</td>
</tr>
<tr>
<td>MR HMM</td>
<td>1.85</td>
<td>3.19</td>
<td>0.58</td>
<td>-0.39</td>
<td>1.84</td>
<td>2.29</td>
<td>0.80</td>
<td>-92</td>
<td>62</td>
</tr>
<tr>
<td>MR Bench</td>
<td>-3.46</td>
<td>3.78</td>
<td>-0.91</td>
<td>-2.71</td>
<td>20.5</td>
<td>3.17</td>
<td>-1.09</td>
<td>-4032</td>
<td>1250</td>
</tr>
<tr>
<td>MSCI Quality</td>
<td>0.50</td>
<td>2.76</td>
<td>0.18</td>
<td>0.20</td>
<td>2.02</td>
<td>1.90</td>
<td>0.26</td>
<td>-208</td>
<td>837</td>
</tr>
<tr>
<td>MSCI Enhanced Value</td>
<td>0.03</td>
<td>3.97</td>
<td>0.01</td>
<td>0.029</td>
<td>0.86</td>
<td>2.83</td>
<td>0.01</td>
<td>-105</td>
<td>599</td>
</tr>
<tr>
<td>MSCI High Dividend Yield</td>
<td>-2.16</td>
<td>3.22</td>
<td>-0.67</td>
<td>0.38</td>
<td>0.85</td>
<td>2.24</td>
<td>-0.96</td>
<td>-2374</td>
<td>1317</td>
</tr>
<tr>
<td>MSCI Momentum</td>
<td>2.48</td>
<td>4.35</td>
<td>0.57</td>
<td>-0.35</td>
<td>1.42</td>
<td>3.11</td>
<td>0.80</td>
<td>-144</td>
<td>475</td>
</tr>
<tr>
<td>MSCI Minimum Volatility</td>
<td>-0.89</td>
<td>3.58</td>
<td>-0.25</td>
<td>0.10</td>
<td>0.69</td>
<td>2.52</td>
<td>-0.35</td>
<td>-38371</td>
<td>906</td>
</tr>
<tr>
<td>MSCI Equal Weighted</td>
<td>-0.27</td>
<td>2.94</td>
<td>-0.09</td>
<td>-0.045</td>
<td>0.74</td>
<td>2.09</td>
<td>-0.13</td>
<td>-135</td>
<td>675</td>
</tr>
<tr>
<td>Model</td>
<td>Ann ret</td>
<td>Ann vol</td>
<td>IR</td>
<td>Skw</td>
<td>kurt</td>
<td>D. risk</td>
<td>Sortino</td>
<td>DD</td>
<td>DD days</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>Sharpe FSHMM</td>
<td>0.061</td>
<td>0.50</td>
<td>0.12</td>
<td>-0.71</td>
<td>2.85</td>
<td>0.37</td>
<td>0.16</td>
<td>-94</td>
<td>387</td>
</tr>
<tr>
<td>Sharpe HMM</td>
<td>-0.11</td>
<td>0.65</td>
<td>-0.16</td>
<td>-0.70</td>
<td>3.84</td>
<td>0.49</td>
<td>-0.22</td>
<td>-164</td>
<td>522</td>
</tr>
<tr>
<td>Sharpe Bench</td>
<td>-1.62</td>
<td>0.92</td>
<td>-1.76</td>
<td>-2.75</td>
<td>15.0</td>
<td>0.82</td>
<td>-1.98</td>
<td>19825</td>
<td>1452</td>
</tr>
<tr>
<td>Dyn FSHMM</td>
<td>0.39</td>
<td>0.65</td>
<td>0.61</td>
<td>-0.41</td>
<td>0.84</td>
<td>0.47</td>
<td>0.84</td>
<td>-52</td>
<td>141</td>
</tr>
<tr>
<td>Dyn HMM</td>
<td>-0.02</td>
<td>0.60</td>
<td>-0.03</td>
<td>-1.12</td>
<td>9.03</td>
<td>0.45</td>
<td>-0.04</td>
<td>-175</td>
<td>566</td>
</tr>
<tr>
<td>Dyn Bench</td>
<td>-1.10</td>
<td>1.03</td>
<td>-1.07</td>
<td>-2.76</td>
<td>16.2</td>
<td>0.88</td>
<td>-1.24</td>
<td>-1508</td>
<td>1123</td>
</tr>
<tr>
<td>MR FSHMM</td>
<td>2.02</td>
<td>3.20</td>
<td>0.63</td>
<td>-0.39</td>
<td>1.83</td>
<td>2.30</td>
<td>0.88</td>
<td>-82</td>
<td>62</td>
</tr>
<tr>
<td>MR HMM</td>
<td>1.85</td>
<td>3.19</td>
<td>0.58</td>
<td>-0.39</td>
<td>1.84</td>
<td>2.29</td>
<td>0.80</td>
<td>-92</td>
<td>62</td>
</tr>
<tr>
<td>MR Bench</td>
<td>-3.46</td>
<td>3.78</td>
<td>-0.91</td>
<td>-2.71</td>
<td>20.5</td>
<td>3.17</td>
<td>-1.09</td>
<td>-4032</td>
<td>1250</td>
</tr>
<tr>
<td>MSCI Quality</td>
<td>0.50</td>
<td>2.76</td>
<td>0.18</td>
<td>0.20</td>
<td>2.02</td>
<td>1.90</td>
<td>0.26</td>
<td>-208</td>
<td>837</td>
</tr>
<tr>
<td>MSCI Enhanced Value</td>
<td>0.03</td>
<td>3.97</td>
<td>0.01</td>
<td>0.029</td>
<td>0.86</td>
<td>2.83</td>
<td>0.01</td>
<td>-105</td>
<td>599</td>
</tr>
<tr>
<td>MSCI High Dividend Yield</td>
<td>-2.16</td>
<td>3.22</td>
<td>-0.67</td>
<td>0.38</td>
<td>0.85</td>
<td>2.24</td>
<td>-0.96</td>
<td>-2374</td>
<td>1317</td>
</tr>
<tr>
<td>MSCI Momentum</td>
<td>2.48</td>
<td>4.35</td>
<td>0.57</td>
<td>-0.35</td>
<td>1.42</td>
<td>3.11</td>
<td>0.80</td>
<td>-144</td>
<td>475</td>
</tr>
<tr>
<td>MSCI Minimum Volatility</td>
<td>-0.89</td>
<td>3.58</td>
<td>-0.25</td>
<td>0.10</td>
<td>0.69</td>
<td>2.52</td>
<td>-0.35</td>
<td>-38371</td>
<td>906</td>
</tr>
<tr>
<td>MSCI Equal Weighted</td>
<td>-0.27</td>
<td>2.94</td>
<td>-0.09</td>
<td>-0.045</td>
<td>0.74</td>
<td>2.09</td>
<td>-0.13</td>
<td>-135</td>
<td>675</td>
</tr>
<tr>
<td>Model</td>
<td>Ann ret</td>
<td>Ann vol</td>
<td>IR</td>
<td>Skw</td>
<td>kurt</td>
<td>D. risk</td>
<td>Sortino</td>
<td>DD</td>
<td>DD days</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Sharpe FSHMM</td>
<td>0.061</td>
<td>0.50</td>
<td>0.12</td>
<td>-0.71</td>
<td>2.85</td>
<td>0.37</td>
<td>0.16</td>
<td>-94</td>
<td>387</td>
</tr>
<tr>
<td>Sharpe HMM</td>
<td>-0.11</td>
<td>0.65</td>
<td>-0.16</td>
<td>-0.70</td>
<td>3.84</td>
<td>0.49</td>
<td>-0.22</td>
<td>-164</td>
<td>522</td>
</tr>
<tr>
<td>Sharpe Bench</td>
<td>-1.62</td>
<td>0.92</td>
<td>-1.76</td>
<td>-2.75</td>
<td>15.0</td>
<td>0.82</td>
<td>-1.98</td>
<td>19825</td>
<td>1452</td>
</tr>
<tr>
<td>Dyn FSHMM</td>
<td>0.39</td>
<td>0.65</td>
<td>0.61</td>
<td>-0.41</td>
<td>0.84</td>
<td>0.47</td>
<td>0.84</td>
<td>-52</td>
<td>141</td>
</tr>
<tr>
<td>Dyn HMM</td>
<td>-0.02</td>
<td>0.60</td>
<td>-0.03</td>
<td>-1.12</td>
<td>9.03</td>
<td>0.45</td>
<td>-0.04</td>
<td>-175</td>
<td>566</td>
</tr>
<tr>
<td>Dyn Bench</td>
<td>-1.10</td>
<td>1.03</td>
<td>-1.07</td>
<td>-2.76</td>
<td>16.2</td>
<td>0.88</td>
<td>-1.24</td>
<td>-1508</td>
<td>1123</td>
</tr>
<tr>
<td>MR FSHMM</td>
<td>2.02</td>
<td>3.20</td>
<td>0.63</td>
<td>-0.39</td>
<td>1.83</td>
<td>2.30</td>
<td>0.88</td>
<td>-82</td>
<td>62</td>
</tr>
<tr>
<td>MR HMM</td>
<td>1.85</td>
<td>3.19</td>
<td>0.58</td>
<td>-0.39</td>
<td>1.84</td>
<td>2.29</td>
<td>0.80</td>
<td>-92</td>
<td>62</td>
</tr>
<tr>
<td>MR Bench</td>
<td>-3.46</td>
<td>3.78</td>
<td>-0.91</td>
<td>-2.71</td>
<td>20.5</td>
<td>3.17</td>
<td>-1.09</td>
<td>-4032</td>
<td>1250</td>
</tr>
<tr>
<td>MSCI Quality</td>
<td>0.50</td>
<td>2.76</td>
<td>0.18</td>
<td>0.20</td>
<td>2.02</td>
<td>1.90</td>
<td>0.26</td>
<td>-208</td>
<td>837</td>
</tr>
<tr>
<td>MSCI Enhanced Value</td>
<td>0.03</td>
<td>3.97</td>
<td>0.01</td>
<td>0.029</td>
<td>0.86</td>
<td>2.83</td>
<td>0.01</td>
<td>-105</td>
<td>599</td>
</tr>
<tr>
<td>MSCI High Dividend Yield</td>
<td>-2.16</td>
<td>3.22</td>
<td>-0.67</td>
<td>0.38</td>
<td>0.85</td>
<td>2.24</td>
<td>-0.96</td>
<td>-2374</td>
<td>1317</td>
</tr>
<tr>
<td>MSCI Momentum</td>
<td>2.48</td>
<td>4.35</td>
<td>0.57</td>
<td>-0.35</td>
<td>1.42</td>
<td>3.11</td>
<td>0.80</td>
<td>-144</td>
<td>475</td>
</tr>
<tr>
<td>MSCI Minimum Volatility</td>
<td>-0.89</td>
<td>3.58</td>
<td>-0.25</td>
<td>0.10</td>
<td>0.69</td>
<td>2.52</td>
<td>-0.35</td>
<td>-38371</td>
<td>906</td>
</tr>
<tr>
<td>MSCI Equal Weighted</td>
<td>-0.27</td>
<td>2.94</td>
<td>-0.09</td>
<td>-0.045</td>
<td>0.74</td>
<td>2.09</td>
<td>-0.13</td>
<td>-135</td>
<td>675</td>
</tr>
</tbody>
</table>
Conclusions

1. Using information from HMMs to construct portfolios improves performance wrt single-regime cases.
Conclusions

1. Using information from HMMs to construct portfolios improves performance wrt single-regime cases.

2. Tested on different kinds of portfolios, improvement is more significant in return-oriented portfolios achieving on average an information ratio of 50% annually in excess of market.
Conclusions

1. Using information from HMMs to construct portfolios improves performance wrt single-regime cases.

2. Tested on different kinds of portfolios, improvement is more significant in return-oriented portfolios achieving on average an information ratio of 50% annually in excess of market.

3. We incorporated embedded feature selection algorithm to our systematic trading framework. This improves model’s accuracy and allows for a more objective approach.
Conclusions

1. Using information from HMMs to construct portfolios improves performance wrt single-regime cases.

2. Tested on different kinds of portfolios, improvement is more significant in return-oriented portfolios achieving on average an information ratio of 50% annually in excess of market.

3. We incorporated embedded feature selection algorithm to our systematic trading framework. This improves model’s accuracy and allows for a more objective approach.

4. We tested both models using MSCI USA enhanced factor indices. Portfolios constructed using feature saliency HMM show a higher performance than the same portfolios constructed using full-feature HMM.
Future Work

1. Include macroeconomic series in the training, where the embedded feature selection could potentially solve the problem of selecting relevant economic series.
Future Work

1. Include macroeconomic series in the training, where the embedded feature selection could potentially solve the problem of selecting relevant economic series.

2. This would be particularly interesting for other asset classes such as fixed income.
Future Work

1. Include macroeconomic series in the training, where the embedded feature selection could potentially solve the problem of selecting relevant economic series.

2. This would be particularly interesting for other asset classes such as fixed income.

3. A drawback of using HMMs is selecting the number of latent states beforehand, we could address this using an infinite HMM.
Thank you!

Paper pre-print: https://arxiv.org/abs/1902.10849
or scan QR code!

https://github.com/elifons/FeatureSaliencyHMM