NUS INVESTMENT SOCIETY

Academic Research Review

Classifying Market Conditions Using Hidden Markov Model

INTRODUCTION

Best known for their applications in speech recognition, Hidden Markov Models (HMMs) are able to discern and recognise underlying patterns through a statistical approach. Parallels between pattern recognition in speech and financial time series naturally lead one to wonder if HMMs can yield returns for a trader. When Renaissance Technologies' Medallion fund was launched in the late 1980s by James Simons, Lenny Baum was one of those hired by Simons to formulate trading strategies. His Baum-Welch algorithm for HMMs eventually became part of the initial strategy that raked in profits for the fund (Patterson, 2011). Existing research has mostly demonstrated the accuracy and usefulness of HMMs at identifying regimes and trend reversals. This paper seek to use HMM in algorithmic trading systems.

KEY ASSUMPTIONS

1. Weak form of market efficiency. This hypothesis dictates that all information about past prices has been fully reflected in the current price. This means we cannot outperform the market by simply analyzing past price data with technical analysis.
2. Markov property. This property states that future price movements are independent of past history, but only dependent on the present price (state) of the market.

MODEL

Hidden Markov Model (HMM) is a probabilistic model to infer true states obfuscated by random observations by analyzing observed time-series data. We name this sequence of observed data O. Correspondingly, we name the sequence of true states H. Note we do not know what these true states are, because each state will randomly produce a range of different observations.

The key parameters of HMM:

N	=	Number of true states in HMM
M	=	Number of possible observations for each state
$\{a_{ij}\}$	=	Transition Probability Matrix of size N x N
a_{ij}	=	Probability of moving from state i to state j
$\{b_j(k)\}$	=	Probability of geting kth observation given state i
π	=	Initial probabilistic condition of HMM

METHODOLOGY

We would like to train the HMM to infer a sequence of true states which most likely can produce the sequence of observations. At the same time, we are able to obtain the associated probability matrices $\{a_{ij}\}$ and $\{b_j(k)\}$. This enables us to predict future observations, i.e. market prices.

In order to estimate parameters of a HMM, we will input observation sequence and number of true states into the following algorithms.

1. Expectation-Maximization (EM) Algorithm. With an initial guess of these parameters, this algorithm helps us to gradually converge to the most likely sequence of true states.

$$\gamma_t(i) = P(H_t = S_i \mid O, \text{parameters})$$

$$\varepsilon_t(i,j) = P(H_t = S_i \text{ and } H_{t+!} = S_j \mid O, \text{parameters})$$

$$\sum_{t=1}^{T-1} \gamma t(\mathrm{i}) \quad = \text{expected number of transitions out of state i during a sequence}$$

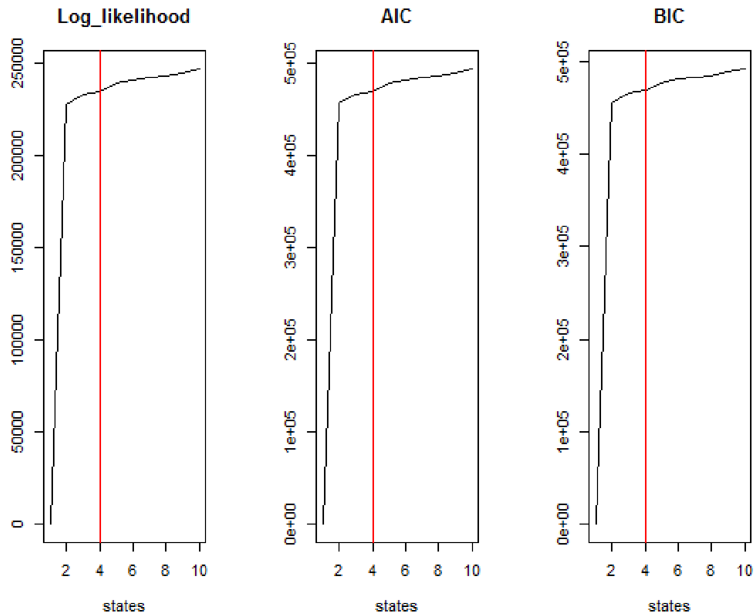
$$\sum_{t=1}^{T-1} \varepsilon t(\mathrm{i},\mathrm{j}) \quad = \text{expected number of transitions out of state i and into state j during a sequence}$$

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \gamma t(\mathrm{i})}{\sum_{t=1}^{T-1} \varepsilon t(\mathrm{i},\mathrm{j})}$$

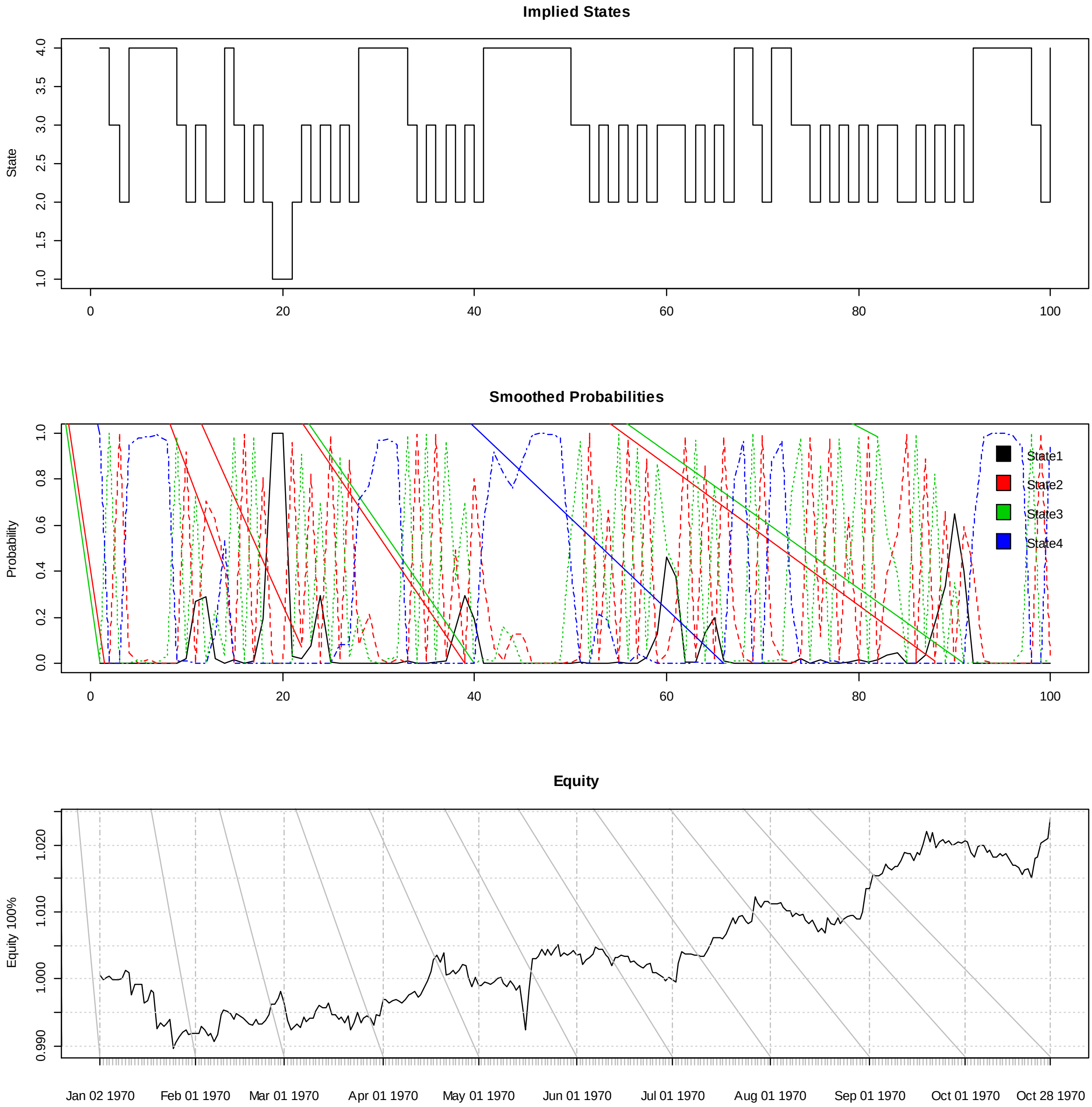
2. Dynamic Programming. This enables us to cut the exponential computation time to O(n^2). One example is Viterbi algorithm which allows us to calculate most probable path given a sequence of observations.

RESULTS

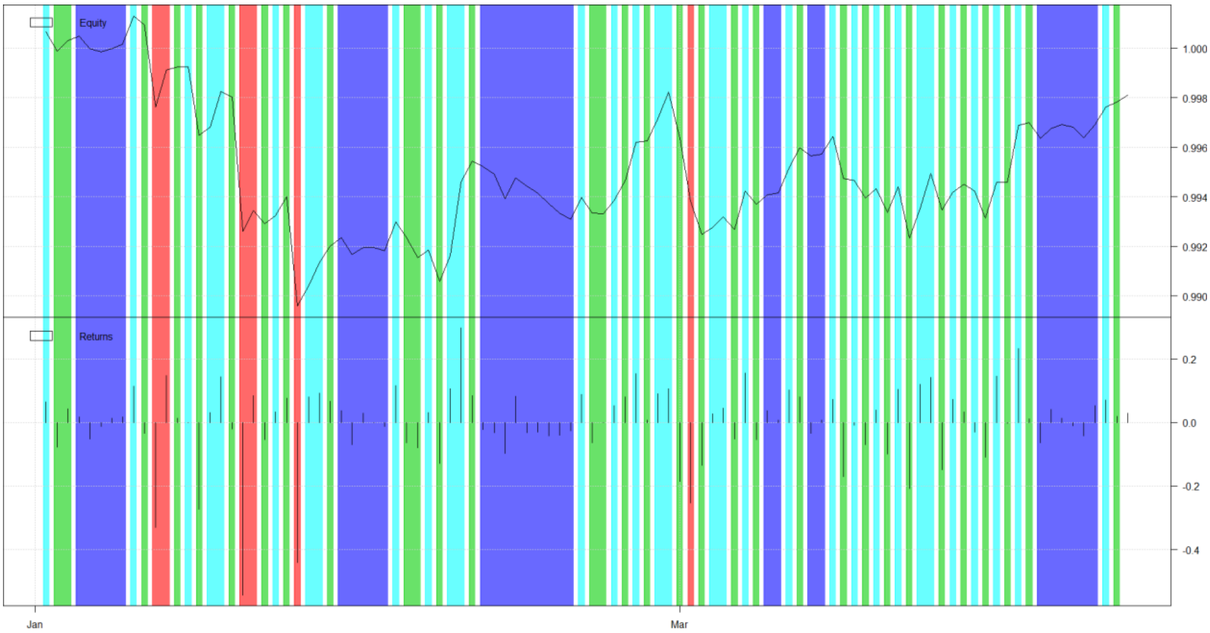
Comparing the evaluative criteria for 2 to 10 number of hidden states, the optimal number of states is 4 which had a least increase in AIC (for small sample size) and BIC (penalize more on parameters count) from 3 states.



As detailed below, State1="Increasing Volatility", State2="Bull", State3="Bear" and State4="Decreasing Volatility". Deduce that high 60% probability of switching between state2 and state3 implies a more unpredictable and whipsaw market with no resting phrase in between. Market stays again in expanding range at 68%. Market stays again in tightening range at 82%. Market rests longer than it jumps out of resting phrase.



Sample of 100 hours of hourly data with smoothed probabilities of being in states 1,2,3,4. State1 can be seen as a trigger for leaving an old range to a new trending phrase. State4 persist at 82% for long periods with an inactive market. Market changes between up and down while trending continuously while trending, indicating a possible strategy of mean reversion while trending and more allowance in trailing stop loss.



When underlying returns, upside variance and downside variance implies a changing normal distribution, it indicates a different state. State1 (red) is a precursor of a new trend. State4 (blue) identifies resting phrase.

```
Transition matrix:
            State 1    State 2    State 3
State 1 0.686218936 0.15672749 0.15705358
State 2 0.063193505 0.19870320 0.61087442
State 3 0.070409273 0.63618365 0.17208553
State 4 0.007710751 0.08392669 0.09110891
             State 4
State 1 1.183152e-09
State 2 1.272289e-01
State 3 1.213215e-01
State 4 8.172536e-01

Conditional distribution parameters:

Distribution parameters:
  State 1
            mean   cov matrix
     -1.185376e-03 0.0734205859 0.0416688969
      5.814471e-03 0.0416688969 0.0704075934
     -5.533496e-05 0.0004068126 0.0002798048

     4.068126e-04
     2.798048e-04
     6.594724e-06

  State 2
            mean   cov matrix
     -5.914795e-02 6.651311e-03 2.934041e-03
     -4.887075e-03 2.934041e-03 4.026163e-03
     -3.056135e-05 7.198686e-05 4.448085e-05

     7.198686e-05
     4.448085e-05
     1.292375e-06

  State 3
           mean   cov matrix
     5.866521e-02 6.488335e-03 7.126196e-03
     1.829217e-03 7.126196e-03 1.325916e-02
     1.275557e-05 2.809382e-05 4.215214e-05

     2.809382e-05
     4.215214e-05
     3.623248e-07

  State 4
            mean   cov matrix
     -7.296414e-05 1.157240e-03 5.999523e-04
     -1.753346e-03 5.999523e-04 1.016835e-03
      1.092001e-05 5.901192e-06 5.776043e-06

     5.901192e-06
     5.776043e-06
     9.561813e-08


Log-likelihood: 236558.4
BIC criterium: -472594.3
AIC criterium: -473014.8
```

Comments:

State 2 and 3 have high interchangeable 2→3 and 3→2 probabilities of 60%. State 1 and State 4 have 68% and 82% probability of staying in their own state in the next period.

State 1: Expanding Range

Returns: N(-0.001%,7%)

Upside Var: Up bias

Downside Var: Down bias

State 2: Bear Market

Returns: N(-0.06%,0.7%)

Upside Var: Down bias

Downside Var: Down bias

State 3: Bull Market

Returns: N(0.06%,0.6%)

Upside Var: Up biases

Downside Var: Up biases

State 4: Tightening Range

Returns: N(-0.00007%,0.1%)

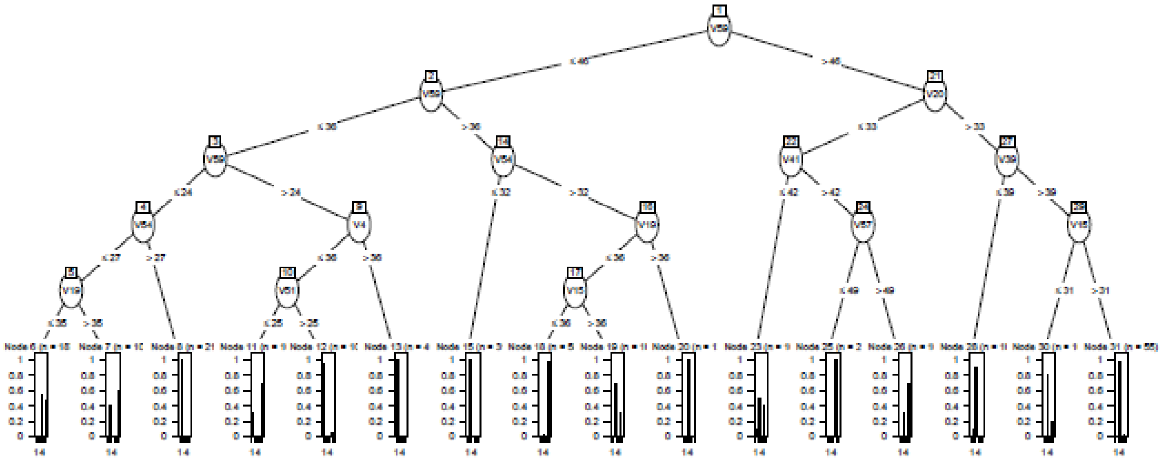
Upside Var: Down biases

Downside Var: Up biases

STRATEGY

A 24/7 market with a tight range long term volatility – as stable currencies create a conducive business environment – with constant influence from daily, weekly and monthly policy meetings and data release. Furthermore, an asset class with the lowest number of tradable assets hence resulting in a highly liquid and hence a “no market fixing“, transparent market – unlike big market cap stocks or commodities traded in a closed network. Pairs are highly correlated as they are dependent on one another as a closed-state triangle relationship. An unpredictable market with instant replenishment of order books and always changing direction, volatility and liquidity. Execution of orders is best left to machines instead of humans. However, machines need a way of learning to make better decisions.

With the least amount of lookback data, conditions for the near immediate future or next 5 time period can be estimated using current conditions: liquidity or book depth, spread, short term actual volatility. Decision tree is used to make effective decisions. Parameters are found using classification tree with Gini index as the measure of classification error, the objective to minimize. Training time series dataset are labeled with implied HMM state as a feature and this experiment is done multiple times. Once classifier is built with confidence, it is used to predict future unlabeled time series data.



Alternating a trend following and mean reversion strategy when state is known. However, more filtering has to be done as time series has a lot of noise. However, this can be counter with fast following from price confirmation without forecasting future short-term price direction. Fast following can be done with proper decision tree of ordering such that the most likely direction is always taken.

Similar to portfolio risk management, trading systems are very asset specific and could be diversified with pairs trading with other correlated and liquid currencies. With risk coming from all directions, safety nets have to be installed. These are the usual procedures of a proper trading systems: well positioned scaling in (pyramiding) and out (partial profiting), post-open and pre-close (unstable spreads due to illiquidity), data release and extreme spreads, max drawdown tolerance, minimum liquidity criteria.

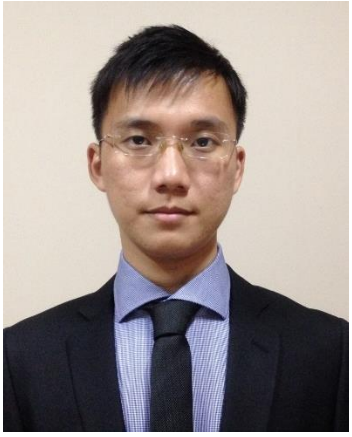
CONCLUSION

While computing power have increased and ability to incorporate numerical models into trading systems is common, latency because of distanced location differences and inefficient algorithm is an issue.

In this example, we only use a discrete HMM. Alternatively, we can use a continuous HMM. Literatures argue that a continuous HMM will produce a better prediction result.

Here we only train our HMM once at the beginning. Actually this training process can continue forever by introducing new observations into the training pool. Also, we can give recent data more weights to focus more on the current trend in our training.

RESEARCH ANALYSTS



Justin Yeo Shui Ming

Zhang Yi Fei

Peter Zhang