Bayesian Hidden Markov Models For Unsupervised Classification of Financial Time Series

Jifeng Sun,Yinghe Qing,Ercan E Kuruoglu
Tsinghua-Berkeley Shenzhen Institute

Abstract

The hidden Markov model is a classical doubly random process, which has been widely used in speech recognition, image analysis and financial timing prediction. In this project, we study the application of the Hidden Markov Model with Gaussian Mixture Distribution in the stock market using Bayesian method like Gibbs sampling for parameter estimation, and compare the performance with the classic Hidden Markov Model using Forward backward algorithm and Viterbi algorithm.

Motivation

- Using MCMC technology to solve optimization problems shows greater advantages than other methods. Comparing with Maximum a Posteriori Estimation like the Viterbi algorithm, the MCMC can give the posterior distributions, while Viterbi only gives a number. It gives a way of error analysis and show how robust about result is.
- In addition, we hope that the method of Hidden Markov model can be used to get a better return on investment. And it may be helpful for the managers make better decisions.

Data Processing

Data acquisition consists of following steps:

- The training dataset is the close price of the zz500 index from 2015-01-06 to 2018-09-07 and the testing dataset is from 2018-09-08 to 2020-06-08.
- the price is processed as three indicators:
  X1: Daily logarithmic rate of return;
  X2: Five-day logarithmic rate of return;
  X3: Daily logarithmic rate of difference.
  The data is fed into the model.

Architecture

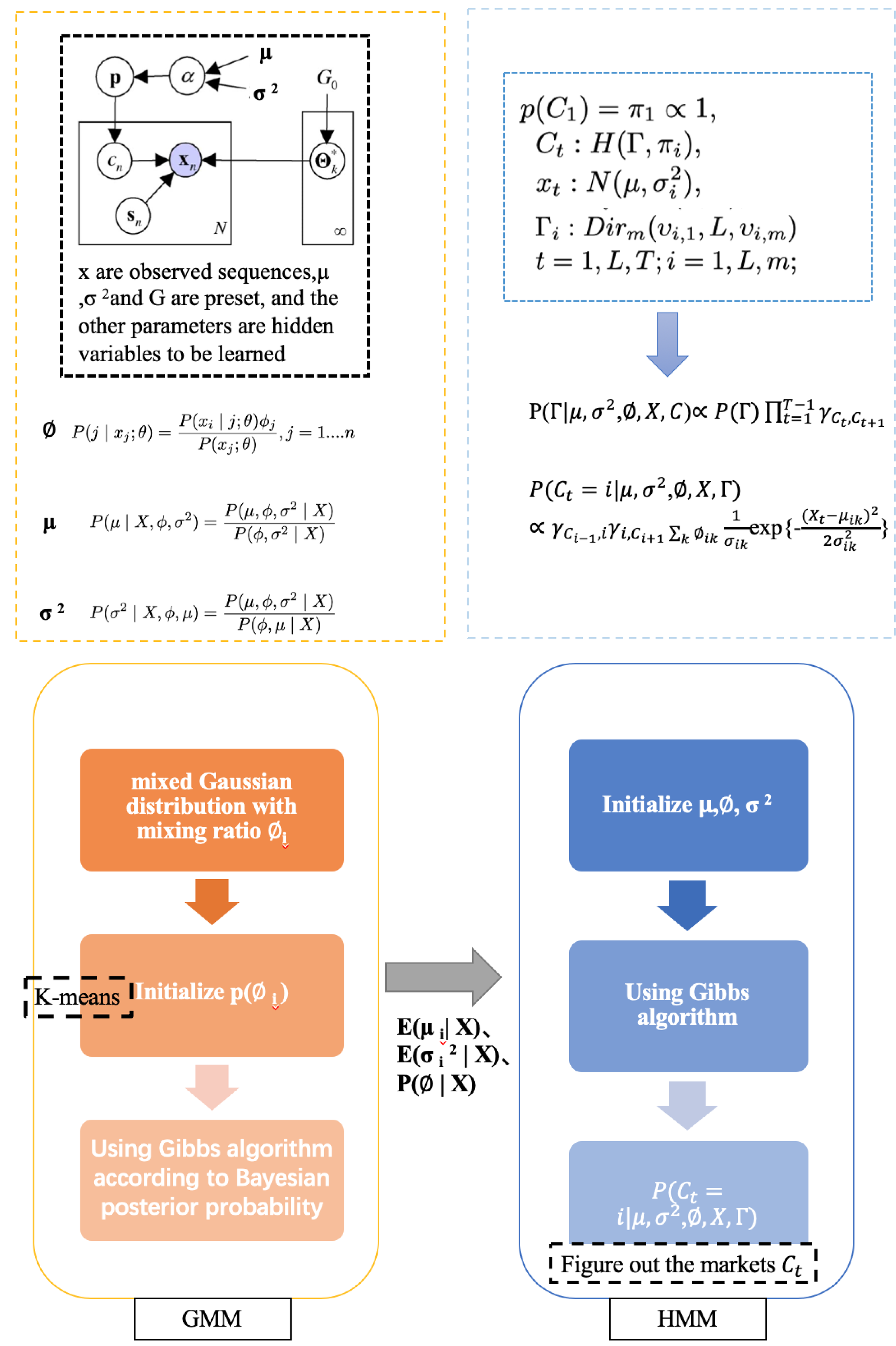

Figure 1:Process flow architecture

Initialize θ_0:$(\theta_{1,0}, ..., \theta_{M,0})$The observation sequence X is$(x_1, ..., x_\tau)$;
for $each\tau \in [1,T]$ **do** ;
Sample$\theta_1^{\tau+1}$ $p(\theta_1|\theta_2^\tau, \theta_3^\tau, ..., \theta_M^\tau, X, M)$;
Sample$\theta_2^{\tau+1}$ $p(\theta_1|\theta_2^\tau, \theta_3^\tau, ..., \theta_M^\tau, X, M)$;
Sample$\theta_i^{\tau+1}$ $p(\theta_1|\theta_1^\tau, \theta_2^\tau, ..., \theta_M^\tau, X, M)$;
Sample$\theta_M^{\tau+1}$ $p(\theta_1|\theta_2^\tau, \theta_3^\tau, ..., \theta_M^\tau, X, M)$;

Figure 2:Gibbs algorithm

Table 1:DBI of three algorithms.The lower the DBI, the better the classification effect.

	bayesGMMHMM	bayesHMM	HMM
DBI	0.9425987	0.9478923	1.353358

Experiment & Results

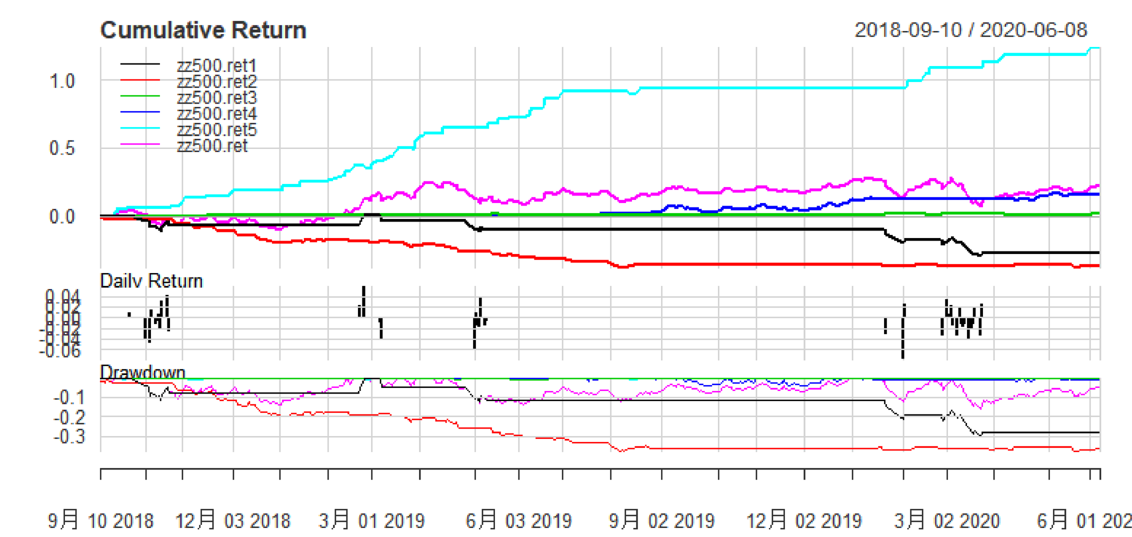

Figure 3:The cumulative return, drawdown of the hidden Markov model using Gaussian mixture model with five states.The blue one get the largest return and it defeats the market during that time.It distinguishes the bull market correctly.

Table 2:Annualized rate of return and annualized Std Dev of the hidden Markov model using Gaussian mixture model with five states

	state1	state2	state3	state4	state5	zz500
Return	-0.2378	-0.2648	-0.0008	0.1119	0.6370	0.0191
Std Dev	0.1833	0.0922	0.0140	0.0650	0.1097	0.2454

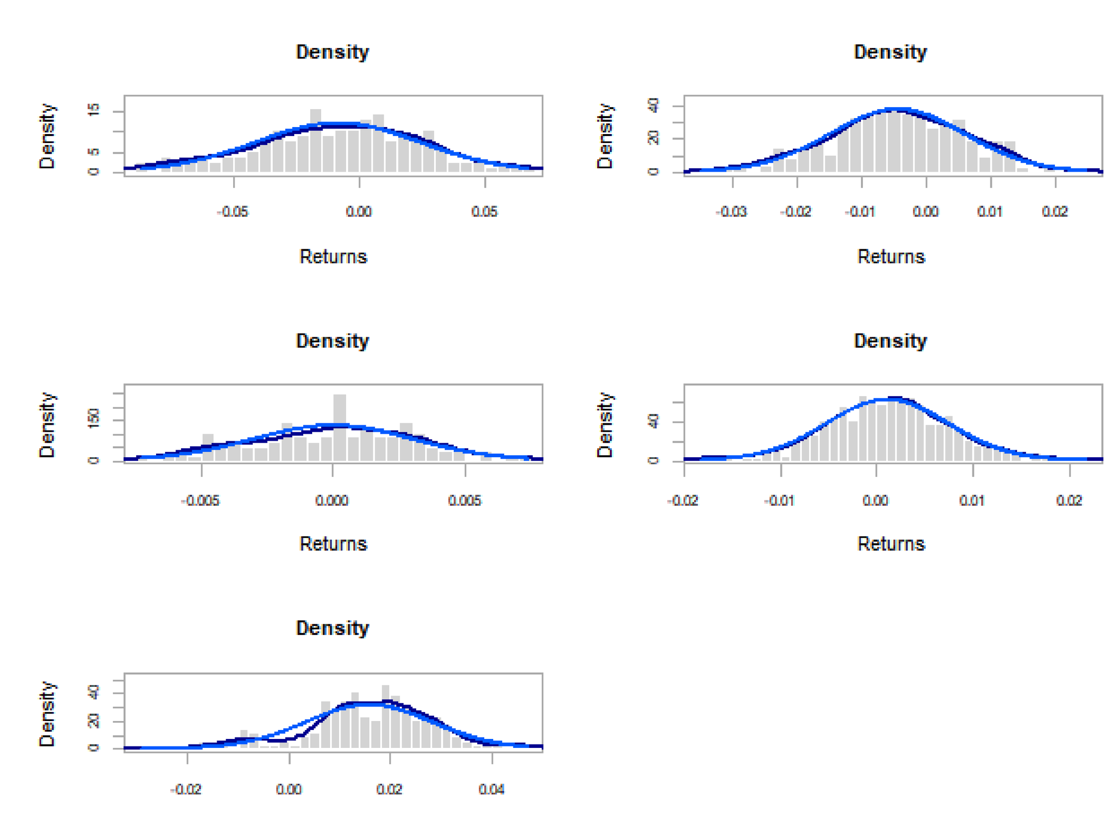

Figure 4:The return's histogram.The light blue line is the probability density curve of a Gaussian distribution with the mean of sample mean and the variance of sample variance.Another line is a fitting curve of the frequency of the sample.

Conclusion & Future Work

In this work, we test the performance of two kinds of Bayesian hidden Markov models on the market index zz500. Their difference is whether the emission distribution is the Gaussian mixture distribution or normal Gaussian distribution. The empirical analysis results show that the mixture Gaussian hidden Markov model effectively estimates the number of states of the model and predicts the rate of return and gets the largest DBI score among three algorithms.

In terms of algorithm improvement, more effective algorithms can be considered, such as the SG-MCMC algorithm. In addition, the classification effect of the algorithm has not been significantly improved. These issues constitute important subjects for future work.

Related Work

- G.D.Forney,"The Viterbi algorithm,"Proceedings of the IEEE,vol.61, no.3, pp.268–278,19
- S.L.Scott, "Bayesian methods for hidden Markov models: Recursive computing in the 21st century,"Journal of the American Statistical Association, vol.97, no.457, pp.337–351,2
- M. Yuting, "Hidden Markov model based on Bayesian analysis and its application," master's thesis, South China Agricultural University,4 201
- H. Jianghua, "Parameter estimation of mixed Gaussian model based on MCMC method,"Journal of Ningxia Normal University, no.3, pp. 10–14,20

Contact Information

- Name: Jifeng Sun, Yinghe Qing
- sjf19@mails.tsinghua.edu.cn, qyh19@mails.tsinghua.edu.cn