

HMM Analysis on Baidu and Alibaba Stock Prices

Jieli Zhou

Description

Alibaba and Baidu are both Chinese internet giants that made their **Initial Public Offering** in NYSE. Baidu went public in 2005 while Alibaba in 2014. They chose New York over Shanghai to go public because China forbids foreign ownership of its companies. However, most of the two Companies' employees, operations and markets are still in China. So they are under the same influence of the bigger Chinese market environment. Though Baidu and Alibaba are of different nature (Baidu provides searching services while Alibaba focuses on E-commerce), their stock prices on NYSE look surprisingly synchronized since Alibaba's IPO. Below is a diagram comparing Alibaba's stock with Baidu's from 2014.



Diagram 1 Stock Price of Alibaba and Baidu Sep 2014 - April 2016

To further show this finding above is very interesting, below is a diagram of stock prices (same period) of two US companies that resemble Baidu and Alibaba and a Chinese social-networking company RenRen.

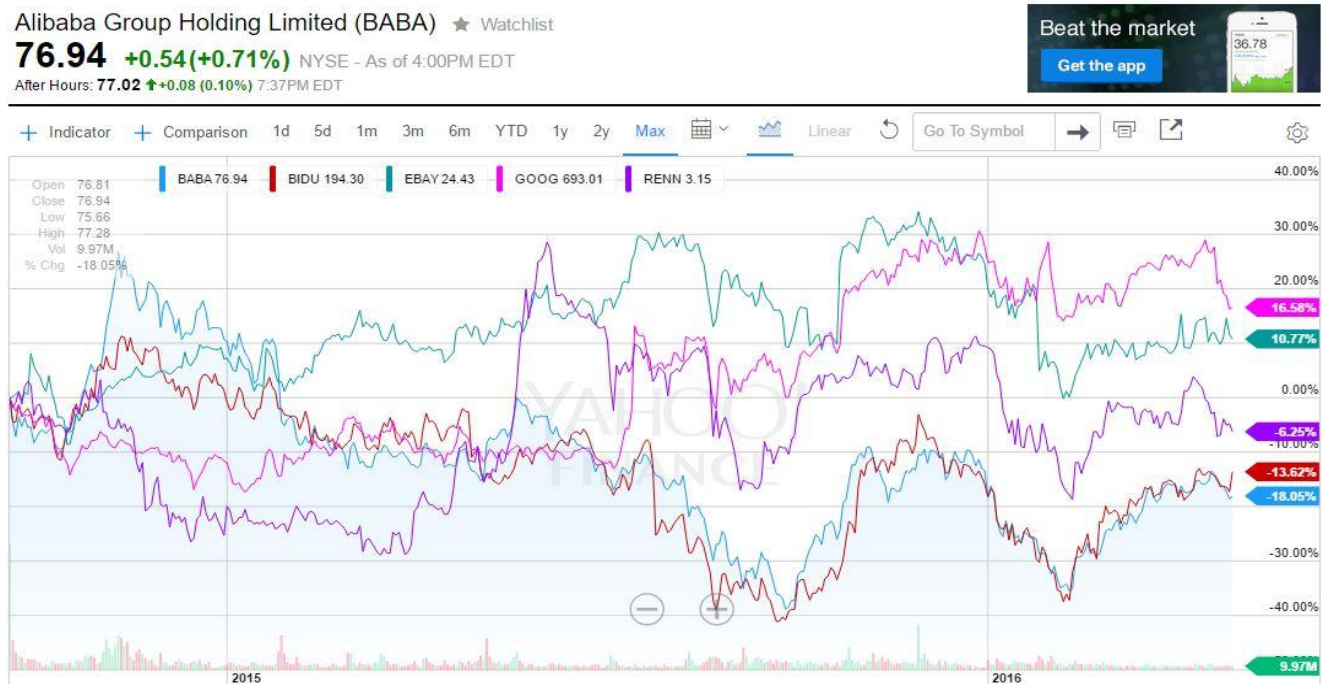


Diagram2 Stock Price of Google Amazon Baidu Alibaba and Renren Sept 2014 – April 2016

It is obvious that other similar companies' stock price are not in-tune as closely as Baidu and Alibaba. So here arise several interesting questions.

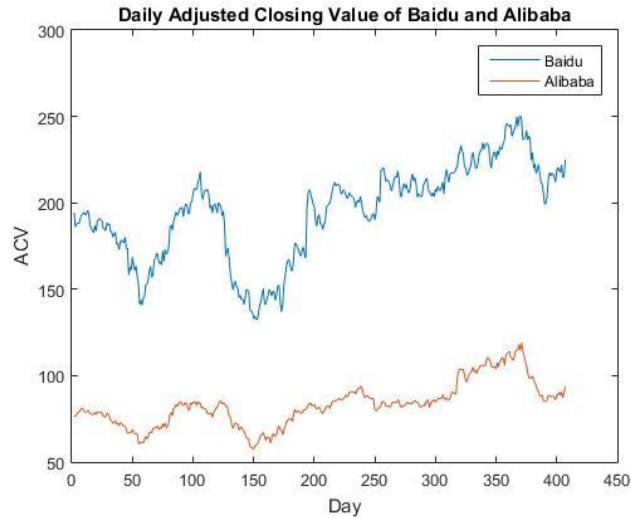
1. What are influencing the up and down of Alibaba's stock price and what are influencing Baidu's?
2. Are there some common influences on both companies, so that their stock prices are in tune?
3. Will they diverge in the future, since Baidu is exploring new field like Artificial Intelligence and Alibaba is expanding its entertainment business.

Due to the limited time constraint, in this project I will focus on the most direct question.

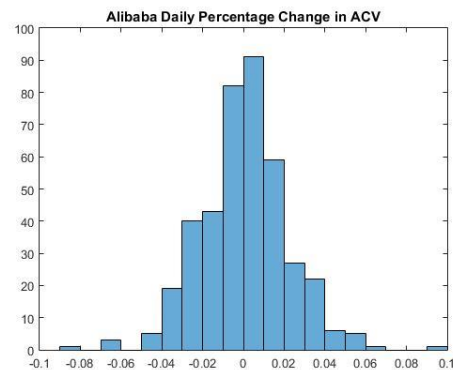
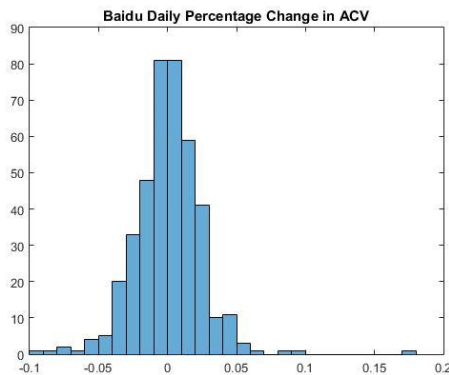
Is Baidu's stock price influencing Alibaba's and Can a HMM be fit?

Data

I used Alibaba and Baidu's stock prices from Sept 2014 to April 2016. I focused on the daily Adjusted Close Prices. Below are two diagrams of the 407 daily adjusted closing prices for both companies



Since Baidu's CV is higher than Alibaba's, I decide to use percentage of change. I compute the daily change of their percentage ACV, and below are the corresponding histograms.



Based on these two histograms, I propose there are 6 hidden states and 6 observation states.

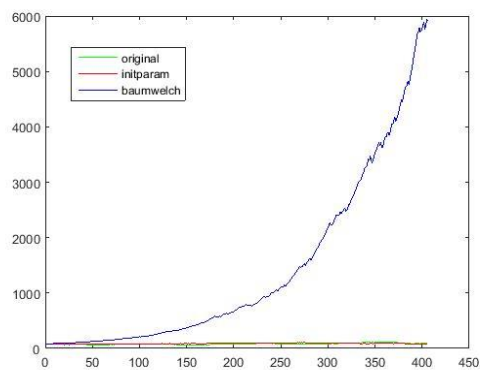
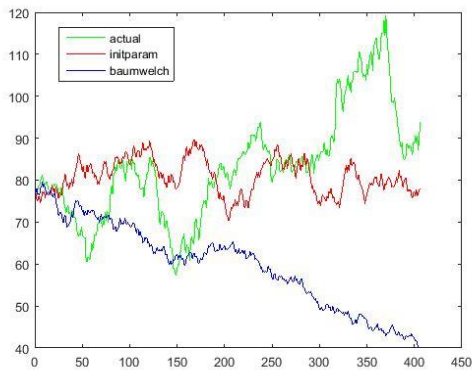
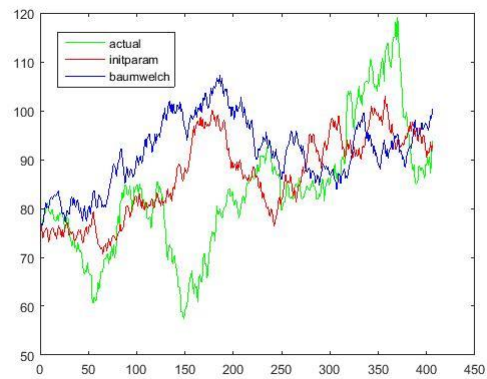
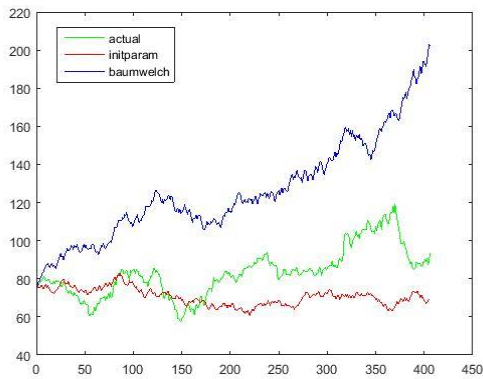
Hidden states are 6 ranges of Baidu's ACV change: -0.1 to -0.02, -0.02 to -0.01, -0.01 to 0, 0 to 0.01, 0.01 to 0.02, 0.02 to 0.2. Observation states are also 6 corresponding ranges of Alibaba's ACV change: -0.1 to -0.02, -0.02 to -0.01, -0.01 to 0, 0 to 0.01, 0.01 to 0.02, 0.02 to 0.2.

Algorithm

1. Propose an initial Transition Matrix and an initial Emission Matrix
2. Use Baum-Welch to update them (Matlab function hmmtrain)
3. Generate a random HMM sequence of Close Values based on the parameters, and plot them together with the actual Alibaba stock price, intuitively see how good the parameters are.

Results

The hmm sequence generated based on the initial T and E (red) is doing ok, but the hmm sequence generated based on the Baum-Welch updated T and E (blue) is very unpredictable.



Conclusion and Further Work

The experiment I did oversimplified the issue. I wanted to show that Baidu's stock price might be a hidden influence on Alibaba's stock price, but the HMM parameter learning process is not good. Clearly, the extent of Baidu's stock fluctuation is not the only hidden influence of Alibaba's, and a batch mode, (considering 407 days together) definitely oversimplified the problem. In addition, a naïve HMM model is not good enough, as literature [1] suggests, continuous HMM, Left-Right (Bakis) HMM and mixture of Gaussian might be more appropriate. The 6 degrees of fluctuations further neglect the actual change.

Doubts I had when I was designing this experiment and reading the literature were How to decide on the number of hidden states, emission states, and what is the 'right' way to do so. This setup in the 2 literature I found confused me, as what exactly is each hidden state doing, are they there only to add to the complexity of the model, or do they actually have meanings. I am clearly layman to stock market, but the literature didn't provide any satisfactory answers as well.

I guess HMM intrigues me because in a clear context, it can provide meaningful predictions, like the wheel-chair paper. However, in a fuzzy context like the stock price prediction, I am totally lost in questions like what hidden states to choose, why I should use Gaussian Mixture and so on.

Though this experiment yields little meaningful results. I think perhaps a HMM analysis on a larger scale could be useful in detecting Financial Crisis and market performance.

References

[1] *Nobakht, Behrooz, Carl-Edward Joseph, and Babak Loni.* "Stock market analysis and prediction using hidden markov models." Student Conference on Engg and Systems (SCES). 2012

[2] *Kavitha G, Udhayakumar A, Nagarajan D.* Stock Market Trend Analysis Using Hidden Markov Models

[3] HMM toolbox, Matlab

[4] Alibaba Baidu Stock Price data, Yahoo Finance