Regime Changes in Non-Stationary Time-Series

Mapleridge Capital Corporation

September 8, 2009

1 Introduction

A stationary process is a stochastic process whose joint probability distribution is invariant under time translation; a stationary time-series is an outcome of a stationary process.

Financial data is certainly not stationary; invariance under time-translation is violated by such effects as volatility-clustering and seasonality.

We wish to understand financial time-series data by modeling it by a multi-regime stochastic process, in which the time-series throughout each regime is generated by a stationary process and the transitions between regimes are so-called 'change points'.

The question which occurs most often when considering regime changes is 'What is the difference between two regimes? What has changed?'. The definition above makes it possible that anything could have changed but let us consider some concrete examples:

- Consider a stock price process, the simplest and most apparent regime change would be a change in the direction of price changes. E.g. the price tended to go up and then it tends to go down.

- Also on the example of the stock price process we might observe a sudden change in the volatility of the returns (we believe this type of regime change is extremely common). More generally we might consider changes in the distribution, incorporating mean, variance and higher moments, of the returns at some point.

- If we believe the processes that generate this sort of financial data is not necessarily an independent return process, we may also detect regime changes where the auto-correlation structure of the returns has changed. More generally we might consider changes in process dynamics, such as a tendency to trend to a tendency to reverse.

- If we consider a basket of stock prices. We might observe regime changes within the correlation matrix of the stock prices.
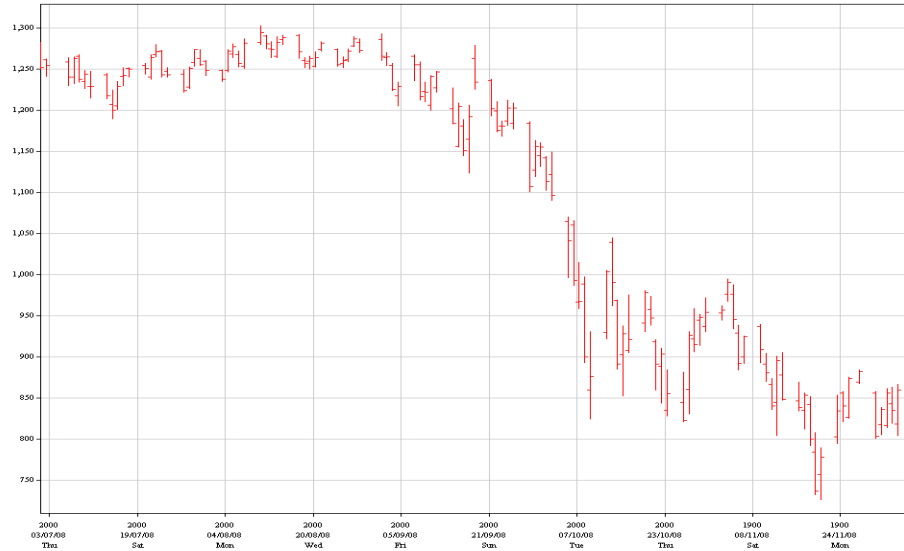
Figure 1: SP 500 Futures Prices during the credit crunch, was there a regime change at this time?

2 Approaches

There are several approaches to this problem of varying sophistication and complexity. The simplest way to detect a single change point in an independent return process would be to measure the likelihood of each point in the time-series being a change point by performing a statistical test (such as the Kolmogorov-Smirnov test) on the data to the left and right of that point. The most likely candidate point could then be accepted or not as the change point.

This method has several drawbacks:

- This can only detect changes in the distribution and will not detect changes in auto-correlations, what tests are available to detect more general changes.

- How do we move on to detecting 2 or more change points when the brute force method of testing each multi-dimensional change point configuration seems intractable? It may be possible to start with detecting a single change point (or at least candidates of such) and then moving on to detecting further change points by analysing each regime. It is not clear, though, that any optimal partitioning of the data with 2 change points, say, is *reducible* to the optimal partioning of the data with 1 change point.

If the number of change points is known beforehand then the most powerful method is to use a hidden Markov model (HMM) to determine the most likely

location of those change points. The topology of the HMM is constrained to only allow 'right-left' transitions, so that the transition matrix is of the form:

$$
\begin{vmatrix}
1 - p_{1,2} & p_{1,2} & 0 & \ldots & 0 \\
0 & 1 - p_{2,3} & p_{2,3} & \ldots & 0 \\
\vdots & & \ddots & & \vdots \\
0 & \ldots & \ldots & 1 - p_{n-1,n} & p_{n-1,n} \\
0 & \ldots & \ldots & 0 & 1
\end{vmatrix} . \tag{1}
$$

The process associated with each state is essentially arbitrary; as long as each can be optimised using maximum likelood, the entire HMM can be optimised using a backwards-forwards (Baum-Welch) algorithm. Once optimised, the HMM and the orignal dataset can be used to obtain a joint probability distribution for the location of the change points. Most practical applications of HMMs use state processes that are iid, most usually a mixed Guassian model, while a simple histogram could be used as a non-parametric model. Yet, there is no reason why this process cannot be an ARMA or other non-iid process. Such would allow modeling of changes in auto-correlation structure.

3   Problems

A major problem is that we do not know how many change points a given sequence has. Simply maximizing the likelihood leads us to a model with as many change points as data points, in which the data set is described exactly but provides us with no understanding of the process or its generalities. We see three approaches to addressing this problem

- A bottom-up approach, in which a single change point is detected and then subsequent change-points are added until no other significant change point can be added. This approach would require a full understanding and resolution of the reducibility problem.

- A top-down approach where each point is assumed to be a change point and neighbouring points are coalesced into regimes until no more combinations seem possible. This approach seems powerful at first, however, it must be understood that the definition of a change point depends on the global properties of the time-series, so any test developed to combine neighbouring points would have to be global in nature.

- An extension of the HMM approach, in which the topology of the transition matrix is 'optimised'. How this is acheived is not entirely clear. Two methods have occurred to us. The first is to introduce some penalty function associated with the topology and then maximize likelihood minus penalty, the penalty could (should?) be inspired by an information-theoretic description. The second approach is to apply a method like simulated annealing in a top-down or bottom-up methodology.

We would like the students to break into three groups, with each group assessing and extending one of the above approaches, including: addressing each approach's specific concerns, furthering the analytic understanding of the problem and working towards practical algorithms. The groups should then share ideas and developments between themselves and work towards a final assessment of the utility, power, drawbacks and pitfalls of the different approaches. Of course, we may be overlooking something in our outline of this problem and if during the workshop a radically different approach is considered we would very much like to hear about it.

References

[1] Rabiner Lawrence R., *A tutorial in hidden Markov models and selected applications in speech recognition*, fellow, IEEE

[2] Fwu, J.-K. and Djuric P.M. *Automatic segmentation of piecewise constant signal by hidden Markov Models*, Proceedings of 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing. **Jun 1996** 283-286

[3] Gales, Mark and Young, Steve, *The theory of segmental hidden Markov models*, Technical Report CUED/F-INFENG/TR Cambridge University Engineering Department (1993) 133

[4] Gereneser, L. Molnar-Saska, G., *Change detection of hidden Markov models*, Decision and Control, 2004. CDC. 43rd IEEE Conforence on **2** (2004) 1754-1758

[5] Hawkins, M., *Point estimation of the parameters of precewise regression models* Applied Statistics **25(1)** (1976) 51-57

[6] Koop, Gary M. and Potter, Simon M. *Forecasting and estimating multiple change point models with an unknown number of change points* Federak Reserve Bank of New York Staff Reports