Indian Institute of Management Kozhikode

Working Paper

IIMK/WPS/172/FIN /2015/08

ROLE OF STYLIZED FEATURES IN CONSTRUCTING ESTIMATORS FOR REGIME SWITCHING MODELS

L Ramprasath

Role of stylized features in constructing estimators for regime switching models

L. Ramprasath
Indian Institute of Management, Kozhikode
Kerala, INDIA 673570
lrprasath@iimk.ac.in

Abstract

This article explores a link between stylized features and estimation accuracy, in the context of estimating the transition probabilities in regime switching models. We provide an example where estimators that are constructed primarily to capture stylized features, need not perform better than the usual estimators. We show this for finite samples, using both simulations and analytical comparisons.

Key words: Regime switching, Stylized feature, Volatility clustering, Estimation, Method of moments, Markov switching

1 Introduction

Many studies have found significant autocovariances in the squared values of economic time series. Timmermann [2000] attributes the success of ARCH models to the fact that these models exhibit such autocovariance patterns. They show that the class of regime switching models also exhibits autocovariances in the squared values. The volatility clustering produced by these models is usually linked to these non-zero autocovariances.

Due to this link, one can use the sample autocovariances of the squared values to construct estimators, for estimating the parameters of a regime switching model, hoping to better capture volatility clustering effects. We will denote them as AIS estimators (estimators based on Autocovariances In the Squares of a time series). The typical justification is: by using such estimators, one can get closer in knowing the true data generating process (for example, Henriksen [2011]). However the question remains as to whether such estimators that are designed to capture stylized features, should always be preferred, whenever they are available. This is a more general question, which can be raised even outside the regime switching context.

Also this is a relatively difficult question, since in most situations, we cannot establish a complete superiority of one estimator over the other, in terms of estimation accuracy. Here we answer the question in negative by providing a counterexample. We show for the regime switching model, the AIS estimators could be inferior to a simpler set of estimators. We compare the AIS estimators with the estimators based on sample autocovariances in the levels of the series (which we denote as AIL estimators). We find that either of these estimators can be used to estimate the transition probabilities in a regime switching model.

This provides us a warning, in situations where one has to choose between a standard estimator (could be moments, least squares or maximum likelihood) and an estimator particularly designed to capture stylized features. In a related article, Kim and White [2004] provides us another warning related to the over dependence on standard estimators, to detect the presence or absence of certain stylized features.

One of the many characteristics of "good" estimators is admissibility, besides unbiasedness, consistency etc. (Lehmann and Casella [1998]). Typically inadmissible estimators are not preferred, since there is a better competing estimator available. In this article, we show that AIS estimator turns out to be inadmissible, under a special case.

Finally, capturing of stylized features can, and should, motivate construction of newer estimators. However, as we show in this article, this cannot be the sole motivation for recommending their use. One still needs to compare them with existing estimators to ensure their superior performance. We use the example of AIS estimators to only highlight this fact and we do not attempt to compare across methodologies like maximum likelihood, method of moments etc. nor recommend estimators based on sample autocovariances.

The article is organized as follows. The following section reports the findings of the Monte Carlo simulations, after introducing the AIS and AIL estimators. Section 3 compares the theoretical mean squared errors of the two estimators under a special case, and section 4 concludes.

2 Background and Numerical Findings

Our study was motivated by an application of regime switching models to option pricing (Henriksen [2011]), where the use of AIS estimators was justified by their ability to capture stylized features. We follow the same model setup as in Henriksen [2011], which is presented below. Under this model, the regime changes are governed by a time homogeneous Markov chain $Y(t)$, which can move between any one of the k regimes $\{1, 2, \ldots, k\}$. The asset dynamics is given by a traditional geometric brownian motion within each regime; for $i = 1, 2, \ldots, k$,

$$dS_i(t) = \alpha_i S_i(t)\, dt + \beta_i S_i(t)\, dB_t$$

where α_i and β_i are the growth and volatility parameters of the i-th regime.

If we let for $i = 1, \ldots, k$,

$$\chi_i^t = \begin{cases} 1 & \text{if } Y(t) = i \\ 0 & \text{otherwise} \end{cases}$$

be the indicator function, indicating the current state of the Markov chain at time t, then the regime switching asset dynamics can be written in a single equation as,

$$dS(t) = S(t) \left(\sum_{i=1}^{k} \chi_i^t \alpha_i dt + \sum_{i=1}^{k} \chi_i^t \beta_i dB_t \right)$$

We do not observe the $Y(t)$ process, while we observe only the asset prices at times, $t = 0, \Delta t, 2\Delta t, \ldots, n\Delta t$, where Δt represents the frequency of sampling. Correspondingly we have the log returns defined as, $X_i = \ln \frac{S(i\Delta t)}{S((i-1)\Delta t)}$, for $i = 1, \ldots, n$. We use lower case to denote the observed log returns, $x_1, \ldots, x_n$, to distinguish them from the corresponding random variables.

We further assume that the regime changes happen only at these discrete time points. Let Y_i denote the state of the Markov chain during the time interval $[(i-1)\Delta t, i\Delta t]$, $i = 1, \ldots, n$, with transition probabilities p_{ij} defined as,

$$p_{ij} = P(Y_m = j \mid Y_{m-1} = i)$$

for all m and $1 \le i, j \le k$. The stationary distribution of the Markov chain is assumed to exist and is given by $\pi = [\pi_1, \ldots, \pi_k]$. If $\boldsymbol{P}$ stands for the transition matrix, then we have

$$\boldsymbol{P}\,\mathbf{1} = \mathbf{1} \quad \text{and} \quad \pi\,\boldsymbol{P} = \pi \tag{1}$$

where $\mathbf{1}$ is a $k \times 1$ vector of ones.

Using Ito's lemma one can verify that for every m, given $Y_m = i$, X_m follows a Normal distribution with mean $(\alpha_i - \frac{\beta_i^2}{2})\Delta t$ and variance $\beta_i^2 \Delta t$. In the rest of the article, we will use $(\mu_i,\, \sigma_i^2)$ in the place of $((\alpha_i - \frac{\beta_i^2}{2})\Delta t,\, \beta_i^2 \Delta t)$, to simplify notation.

Parameter Estimation

Estimating the parameters in this model amounts to estimating $\pi, \boldsymbol{P}$ and (μ_i, σ_i) for $i = 1, \ldots, k$. It should be noted that between π and $\boldsymbol{P}$, there are some implicit relations arising out of (1). We restrict our analysis to a simpler two regime situation, since with more than two regimes, we estimate multiple transition probabilities simultaneously and we do not have a single mean squared error (MSE) to compare. So we analyze the two regime case, where the comparison is more transparent.

Under a two regime model, we need to estimate only $\mu_1, \sigma_1, \mu_2,\, \sigma_2, \pi_1$ and p_{11}, because of (1). Henriksen [2011] carries out this estimation in two stages; in the first stage they estimate μ_1, σ_1, μ_2, σ_2 and π_1 based on maximum likelihood, and in the second stage, they estimate the transition probabilities by matching autocovariances. For the two regime model, Timmermann [2000] shows,

$$E[(X_i^2 - E[X_i^2])(X_{i-1}^2 - E[X_{i-1}^2])] = \pi_1(1-\pi_1)(\mu_2^2 - \mu_1^2 + \sigma_2^2 - \sigma_1^2)^2(p_{11} + p_{22} - 1)$$

assuming that the underlying Markov chain is starting from its stationary distribution. Using this expression and the estimated values from the first stage, Henriksen [2011] proposes,

$$\hat{\theta}_1 = \frac{1}{(n-1)\pi_1(1-\pi_1)(\mu_2^2 - \mu_1^2 + \sigma_2^2 - \sigma_1^2)^2} \sum_{i=2}^{n} (X_i^2 - E[X_i^2])(X_{i-1}^2 - E[X_{i-1}^2])$$

as the method of moments estimator for $p_{11} + p_{22} - 1$ (the estimated values from the first stage are plugged in the theoretical value of $E[X_i^2]$). This will be our AIS estimator. The estimate of p_{11} can then be obtained using (1). Autocovariances of higher lags are matched when there are more regimes and hence more transition probabilities to be estimated.

Now Timmermann [2000] also shows that for the same model,

$$E[(X_i - E[X_i])(X_{i-1} - E[X_{i-1}])] = \pi_1(1-\pi_1)(\mu_2 - \mu_1)^2(p_{11} + p_{22} - 1)$$

This suggests

$$\hat{\theta}_2 = \frac{1}{(n-1)\pi_1(1-\pi_1)(\mu_2 - \mu_1)^2} \sum_{i=2}^{n} (X_i - E[X_i])(X_{i-1} - E[X_{i-1}])$$

can be another method of moments estimator for $p_{11} + p_{22} - 1$ (the AIL estimator). We will now compare the performance of the AIS and AIL estimators for this estimation problem.

Since it is hard to write in closed form, the exact expression for the MSE of either of these estimators, we compare them using a Monte Carlo study. We make the simplifying assumption that all the parameters except the transition probabilities are known. Then both the above estimators are actually unbiased estimators; so clearly the one with the lesser variance (which is same as the MSE because of no bias) is typically preferred. This assumption is reasonable in the setting of Henriksen [2011]'s two stage procedure, where all the other other parameters are estimated in the first stage and only the transition probabilities are estimated in the second stage.

We simulate the regime switching dynamics under various parameter combinations and compute the mean squared error between the true and the estimated transition probability over many samples. One particular parameter combination (the middle row of the table) is discussed in more detail in the next section. We choose the other combinations (of μ_1, σ_1, μ_2, σ_2, and π_1) around this special case.

In Table 1, we report the difference: $\text{MSE}(\hat{\theta}_1) - \text{MSE}(\hat{\theta}_2)$. The true parameter values refer to the values of $p_{11} + p_{22} - 1$. All the entries being positive indicates, that the AIL estimator performs better in terms of MSE. The standard errors for each of the entries, were found to be low enough to make these positive entries statistically significant.

Each entry in the table was obtained using 1000 simulations, where in each simulation a time series of length 250 was simulated under the regime switching dynamics. We notice enough evidence in the two regime case to conclude that AIS has higher variance compared to AIL, thereby indicating that capturing stylized features may not have a direct effect on improving estimation accuracy.

In one of the parameter combinations, where $\mu_1 = 1$, $\sigma_1^2 = 1$, $\mu_2 = 6$ and $\sigma_2^2 = 2$, the two populations are quite far apart and looking at the X_i s, it is relatively easy to label them with their corresponding regimes because of this distance. Thus in this case the transition probabilities across regimes should be easier to estimate given the data. Both AIL and AIS estimators also performed well in this case, as we observed with their individual MSEs.

Table 1: Reduction in simulated mean squared errors when $\hat{\theta}_1$ is replaced by $\hat{\theta}_2$

(a) $\pi_1 = 0.3$

	True parameter values				
	-0.8	-0.2	0	0.2	0.8
$(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$					
(1, 1.2, 1, 1)	#	69.4	67.5	69.7	77.9
(1, -1.2, 1, 1)	#	134.3	129.2	143.6	140.9
(1, 6, 1, 2)	#	0.0056	0.0063	0.0049	0.0048

(b) $\pi_1 = 0.5$

	True parameter values				
	-0.8	-0.2	0	0.2	0.8
$(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$					
(1, 1.2, 1, 1)	42.0	40.6	45.8	42.7	38.0
(1, -1.2, 1, 1)	85.4	78.9	77.2	81.9	85.0
(1, 6, 1, 2)	0.003	0.0025	0.0027	0.0024	0.004

(c) $\pi_1 = 0.7$

	True parameter values				
	-0.8	-0.2	0	0.2	0.8
$(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$					
(1, 1.2, 1, 1)	#	47.2	47.9	45.4	48.2
(1, -1.2, 1, 1)	#	101.8	96.7	101.6	102.7
(1, 6, 1, 2)	#	0.0004	0.0013	0.0022	0.0061

#: These parameter combinations are not possible due to restrictions placed by equation (1)

3 Analytical Comparison

We will analytically compare the variance of AIS and AIL under a special case, which will explain some of the positive entries observed in Table 1. As in section 2, we assume that all parameters other than the transition probability are known and we will compare $Var(\hat{\theta}_1)$ with $Var(\hat{\theta}_2)$. We make these comparisons without actually computing these two variances, as we will show below.

Since they are both unbiased for $p_{11}+p_{22}-1$, it is enough to compare $E(\hat{\theta}_1^2)$ and $E(\hat{\theta}_2^2)$. We note that for all i, $E[X_i] = \pi_1\mu_1 + \pi_2\mu_2$, because of stationarity. We will denote this by ν_1. Similarly we define, $\nu_2 = E[X_i^2] = \pi_1(\mu_1^2 + \sigma_1^2) + \pi_2(\mu_2^2 + \sigma_2^2)$. Consider $E(\hat{\theta}_2^2)$ first.

$$E(\hat{\theta}_2^2) = \frac{1}{\left((n-1)\pi_1(1-\pi_1)(\mu_2-\mu_1)^2\right)^2} E\left(\sum_{i=2}^{n}(X_i-\nu_1)(X_{i-1}-\nu_1)\right)^2.$$

The expectation on the right hand side of the above expression will be a linear combination of three simpler expected values, which are,

$$E\big[(X_i-\nu_1)(X_{i-1}-\nu_1)(X_{i+j}-\nu_1)(X_{i+j-1}-\nu_1)\big] \tag{2}$$
$$E\big[(X_{i-1}-\nu_1)(X_i-\nu_1)^2(X_{i+1}-\nu_1)\big] \tag{3}$$
$$E\big[(X_i-\nu_1)^2(X_{i-1}-\nu_1)^2\big]. \tag{4}$$

For the AIS estimator,

$$E(\hat{\theta}_1^2)=\frac{1}{\big((n-1)\pi_1(1-\pi_1)(\mu_2^2-\mu_1^2+\sigma_2^2-\sigma_1^2)^2\big)^2}E\left(\sum_{i=2}^{n}(X_i^2-\nu_2)(X_{i-1}^2-\nu_2)\right)^2.$$

As with the AIL case, there will be three simpler expected values for the expectation on the right hand side;

$$E\big[(X_i^2-\nu_2)(X_{i-1}^2-\nu_2)(X_{i+j}^2-\nu_2)(X_{i+j-1}^2-\nu_2)\big] \tag{5}$$
$$E\big[(X_{i-1}^2-\nu_2)(X_i^2-\nu_2)^2(X_{i+1}^2-\nu_2)\big] \tag{6}$$
$$E\big[(X_i^2-\nu_2)^2(X_{i-1}^2-\nu_2)^2\big]. \tag{7}$$

Each of these six expected values will be easy to compute and compare, once we condition on the underlying Markov chain. An appendix contains the proof that the conditional expectations of each of (2), (3) and (4) are lesser than the conditional expectations of (5), (6) and (7) respectively, under a special case when $\pi_1=0.5$, $\sigma_1=\sigma_2$ and μ_1 is close to $-\mu_2$. This comparison is done after taking in to account the different scaling constants that appear in $\hat{\theta}_1$ and $\hat{\theta}_2$.

The inequalities for conditional expectations also carries over to the unconditional expectations, since the coefficients of the conditional expectations are the marginal probabilities of the Markov chain, which are same for both $E(\hat{\theta}_1^2)$ and $E(\hat{\theta}_2^2)$. Thus in this special case (which is similar to row 5 in the table of section 2), $\hat{\theta}_1$ has higher variance compared to $\hat{\theta}_2$, for all possible values of $p_{11}+p_{22}-1$.

4 Discussion

We have shown that estimators which are specifically constructed to capture stylized features do not always guarantee better estimation accuracy. They still have to pass the conventional tests for "good" estimators. One can similarly ask 'Whether an accurate estimator will do a good job in capturing stylized features?'. Our answer is: Yes, it does, since essentially we are doing a model fitting exercise, and smaller is the gap between the true and the estimated model, better is the fit. Accurate estimators automatically take care of accurately capturing the stylized features that are implied by the model. Thus one need not worry about the false dichotomy between these two criterion (capturing stylized features and estimation accuracy) in choosing an estimating procedure. We have also shown that, for a special case, AIS estimators are indeed inadmissible for estimating the transition probabilities of the underlying Markov chain. It will be interesting to see whether this inadmissibility generalizes to AIS based estimators for models with more regimes.

In addition, from our simulations we notice that the improvement of the AIL estimator is higher, when there is a larger overlap between the underlying populations, which is when estimation accuracy plays a more important role. The finite sample MSE comparisons in section 3, are in general difficult to obtain for estimators in models with more number of regimes, due to the presence of covariance terms. We hope that some of these calculations will provide us a useful starting point, for finding improvements to the overall estimation procedure for these models.

Acknowledgement

I would like to thank John Kolassa, Shubhasis Dey and Kausik Gangopadhyay for many valuable comments and suggestions, which improved the presentation.

References

Henriksen, P.N. "Pricing barrier options by a regime switching model." Quantitative Finance, 11(2011), pp. 1221–1231.

Kim, T-H., White, H. "On more robust estimation of skewness and kurtosis." Finance Research Letters, 1(2004), pp. 56–73.

Lehmann, E.L., Casella, G., Theory of Point Estimation, 2nd ed. New York: Springer, 1998.

Timmermann, Allan. "Moments of Markov switching models." Journal of Econometrics, 96(2000), pp. 75–111.

Appendix

We give here the details of the argument that Variance of $\hat{\theta}_2$ is less compared to that of $\hat{\theta}_1$, when $\pi_1 = 0.5$, $\sigma_1 = \sigma_2 (= \sigma)$ and μ_1 is close to $-\mu_2$. We start with the comparison between the expected values (2) and (5). We note that,

$$E[(X_i - \nu_1) \mid Y_i] = \begin{cases} -\left(\frac{\mu_2 - \mu_1}{2}\right) & \text{if } Y_i = 1 \\ \frac{\mu_2 - \mu_1}{2} & \text{if } Y_i = 2 \end{cases}$$

and

$$E(X_i^2 - \nu_2) = \begin{cases} -\left(\frac{\mu_2^2 - \mu_1^2}{2}\right) & \text{if } Y_i = 1 \\ \frac{\mu_2^2 - \mu_1^2}{2} & \text{if } Y_i = 2. \end{cases}$$

Also (2) can be written as,

$$E\big[(X_i - \nu_1)(X_{i-1} - \nu_1)(X_{i+j} - \nu_1)(X_{i+j-1} - \nu_1)\big] = \\ E\big[E\big[(X_i - \nu_1)(X_{i-1} - \nu_1)(X_{i+j} - \nu_1)(X_{i+j-1} - \nu_1)\big] \mid Y_i, Y_{i-1}, Y_{i+j}, Y_{i+j-1}.\big]$$

Similarly for (5).

If we assume that $\mu_1 < \mu_2$ (without loss of generality), $\mu_1^2 < \mu_2^2$ and fix the values for $Y_i, Y_{i-1}, Y_{i+j}, Y_{i+j-1}$, then the conditional expectation for (2) will be exactly same as the conditional expectation for (5), after dividing by the scaling constants that appear in the expressions of $\hat{\theta}_1$ and $\hat{\theta}_2$. We repeatedly use the fact that in this model, the dependence of X_i s is through the Y_i s and once the values of Y_i s are known, the X_i s become independent.

Since all the conditional expectations are equal, this implies that the unconditional expectation in (2), when divided by $(\mu_2 - \mu_1)^4$ and the unconditional expectation in (5), when divided by $(\mu_2^2 - \mu_1^2)^4$ are one and the same. Similar arguments can also be used to show the equality of (2) and (5), when $\sigma_1 \neq \sigma_2$, but this relaxation will make the proof difficult for the remaining two inequalities derived below. So we retain the assumption of $\sigma_1 = \sigma_2$ in what follows.

Next we compare (3) and (6). For a two-state Markov chain, (Y_{i-1}, Y_i, Y_{i+1}) will have 8 possible values. Like in (2) and (5), the conditional expectations for (3) and (6) will have the same absolute values; only their signs will differ. Combining these terms with the marginal probabilities of the Markov chain, we get,

$$E\big[(X_{i-1}-\nu_1)(X_i-\nu_1)^2(X_{i+1}-\nu_1)\big] = 0.25(\sigma^2+0.25(\mu_2-\mu_1)^2)(\mu_2-\mu_1)^2(2p_{11}-1)^2$$
$$E\big[(X_{i-1}^2-\nu_2)(X_i^2-\nu_2)^2(X_{i+1}^2-\nu_2)\big] = 0.25(2\sigma^4+2\sigma^2(\mu_1^2+\mu_2^2)+0.25(\mu_2^2-\mu_1^2)^2)(\mu_2^2-\mu_1^2)^2(2p_{11}-1)^2$$

We also observe that,

$$\frac{\sigma^2}{(\mu_2-\mu_1)^2} \le \frac{2\sigma^4+2\sigma^2(\mu_1^2+\mu_2^2)}{(\mu_2^2-\mu_1^2)^2},$$

since $2(\mu_1^2+\mu_2^2) \ge (\mu_1+\mu_2)^2$. This shows that the expectation in (3), when divided by $(\mu_2-\mu_1)^4$ is always less than or equal to the expectation in (6) divided by $(\mu_2^2-\mu_1^2)^4$.

Coming to the last comparison between (4) and (7) we have,

$$E\big[E\big[(X_i-\nu_1)^2(X_{i-1}-\nu_1)^2\big] \mid Y_{i-1}=\ell, Y_i=m\big] = (\sigma^2+0.25(\mu_2-\mu_1)^2)^2$$

and

$$E\big[E\big[(X_i^2-\nu_2)^2(X_{i-1}^2-\nu_2)^2\big] \mid Y_{i-1}=\ell, Y_i=m\big] = (2\sigma^4+4\sigma^2\mu_\ell^2+0.25(\mu_2^2-\mu_1^2)^2)(2\sigma^4+4\sigma^2\mu_m^2+0.25(\mu_2^2-\mu_1^2)^2)$$

for $\ell, m \in \{1,2\}$.

Also, $(2\sigma^2+4\mu_m^2)(\mu_2-\mu_1)^2 \ge (\mu_2^2-\mu_1^2)^2$ for $m=1,2$, whenever μ_1 is close to $-\mu_2$ (in such a case the right hand side of the inequality is close to zero). This implies,

$$\frac{\sigma^2}{(\mu_2-\mu_1)^2} \le \frac{2\sigma^4+4\sigma^2\mu_m^2}{(\mu_2^2-\mu_1^2)^2},$$

for $m=1,2$. And this in turn implies that the expectation in (4), when divided by $(\mu_2-\mu_1)^4$ is always less than or equal to the expectation in (7) divided by $(\mu_2^2-\mu_1^2)^4$.

This completes the proof that Variance of $\hat{\theta}_2$ is less compared to that of $\hat{\theta}_1$, under the special case considered here.

Indian Institute of Management Kozhikode

Type of Document: (**Working Paper/Case/ Teaching Note, etc.)** ***Working Paper***	*Ref. No.: (to be filled by RCP office)* **IIMK/WPS/172/FIN /2015/08**
Title: **Role of Stylized features in constructing estimators for regime switching models**	
Author(s):	*Institution(s)*
L Ramprasath	Associate Professor , Indian Institute of Management Kozhikode campus P.O. Kunnamanagalam- 673570 Email: lrprasath@iimk.ac.in
Subject Areas : **Finance, Statistics**	*Subject Classification Codes, if any:*
Supporting Agencies, if any:	*Research Grant/Project No.(s):*
Supplementary Information, if any:	*Date of Issue: (to be filled by RCP office)* **February 2015**
Full text or only abstract to be uploaded on website: (please choose one) ***Full text***	*Number of Pages: 9*

Abstract:

This article explores a link between stylized features and estimation accuracy, in the context of estimating the transition probabilities in regime switching models. We provide an example where estimators that are constructed primarily to capture stylized features, need not perform better than the usual estimators. We show this for finite samples, using both simulations and analytical comparisons.

Key Words/Phrases: Regime switching, Stylized feature, Volatility clustering, Estimation, Method of moments, Markov switching

Referencing Style Followed: ***APA***

Research, Conference And Publication Office
Indian Institute Of Management Kozhikode
IIMK Campus P.O., Kozhikode 673 570
Kerala, India
Telephone +91 495 2809 238
E-mail rcp@iimk.ac.in
website www.iimk.ac.in