Change points and structural breaks

Kevin Kotzé

Contents

Introduction

- Time series variables may be influenced by events that arise at various points in time that alter the underlying data generating process
- This would influence the results of many linear time series models that do not account for such features
- In recent periods of time the effects of the Global Financial Crisis and the Covid-19 pandemic may have given rise to a number of structural breaks that may need to be identified
- Extensive literature that includes recent advances at a level of generality that allow a host of interesting practical applications
- Include models with stationary regressors and errors that can exhibit temporal dependence and heteroskedasticity, models with trending variables and possible unit roots and cointegrated models, among others

Introduction

- Procedures may be used to test for common breaks across a large number of variables in either the mean or the variance of a time series or a component thereof, as well as changes in forecast accuracy
- Our focus is on specific aspects that relate to econometric applications that are based on linear relationships between variables
- Make use of retrospective (or *offline*) methods that test for breaks in a given sample of data and form confidence intervals around the break dates
- For recent reviews of the literature, see Perron (2006), Perron (2010), Casini & Perron (2019)

Change point detection

- Change point test seeks to identify the specific period of time that relates to a change in the probability distribution of a stochastic process
- May need to determine whether or not a change has occurred or whether several changes might have occurred
- Assume that we have an ordered sequence of data, $y_t = \{y_1, \ldots, y_T\}$
- Change point is then said to arise when there exists a time, $\tau \in \{1, \ldots, T-1\}$, such that the statistical properties of $\{y_1, \ldots, y_\tau\}$ and $\{y_{\tau+1}, \ldots, y_T\}$ are different
- Could allow for m possible change points that are associated with positions, $\boldsymbol{\tau}_{1:m} = \{\tau_1, \ldots, \tau_m\}$
- Each change point position is then ordered so that $\tau_i < \tau_j$ for $i < j$
- The parameters associated with each segment may be denoted $\{\theta_i, \phi_i\}$, where ϕ_i is a possible set of nuisance parameters and θ_i is the set of parameters that may describe the change

Single change point detection

- To consider the identification of a single change point we could use a likelihood based framework
- Construct a hypothesis test, where the null hypothesis, H_0, suggest that are no change points, $(m = 0)$, and the alternative hypothesis, H_1, would suggest that there is a single change point, $(m = 1)$
- Hinkley (1970) has then shown that we can derive an asymptotic distribution for the likelihood ratio test statistic for a change in the mean within normally distributed observations
- Gupta & Tang (1987) constructs a similar asymptotic distribution for the change in the variance

Single change point detection

- A likelihood ratio test involves comparing the maximum log-likelihood under both null and alternative hypotheses
- For the null hypothesis the maximum log-likelihood is $\log p(y_{1:T}|\hat{\theta})$, where $p(\cdot)$ is the probability density function for the data and $\hat{\theta}$ are the estimates for the parameters
- For the alternative hypothesis the maximum log-likelihood for a given change point at τ_1 is

$$ML(\tau_1) = \log p(y_{1:\tau}|\hat{\theta}_1) + \log p(y_{\tau+1:T}|\hat{\theta}_2)$$

- Involves the use of the maximum value after considering all the possible positions for τ_1

Single change point detection

- The test statistic is thus

$$\lambda = 2\left[\max_{\tau_1} ML\left(\tau_1\right) - \log p\left(y_{1:T}|\hat{\theta}\right)\right]$$

- Need to choose a threshold, c, such that we reject the null hypothesis if $\lambda > c$
- If we reject the null hypothesis, then value of τ_1 that maximises $ML(\tau_1)$ could be used to provide an estimate of the change point position
- The appropriate value for the threshold parameter, c, is an open research question

Multiple change point detection

- The single change point likelihood test statistic can be extended to multiple changes by considering the log-likelihoods for each of the possible m segments
- Involves the identification of the maximum of $ML(\tau_1)$ over all possible combinations of τ_1 and m
- Most common approach looks to minimise the following expression:

$$\sum_{i=1}^{m+1} \left[\mathcal{C}\left(y_{(\tau_{i-1}+1):\tau_1}\right)\right] + \beta f(m)$$

- where $\mathcal{C}$ is a cost function for a segment and $\beta f(m)$ is a penalty function that could take the form of information criteria, such as the AIC, BIC, etc.

Multiple change point detection

- Binary segmentation is arguably the most widely used multiple change point search method
- We initially apply a single change point test to all the available data, if a change point is identified the data is split into two at the change point location
- The single change point procedure is then repeated on the two new sub-samples of data
- If additional change points are identified in either of these, they are split further
- Binary segmentation is an approximate algorithm that is computationally expedient as it only considers a subset of the 2^{T-1} possible solutions

Multiple change point detection

- Segment neighborhood algorithm minimises the expression with the aid of a dynamic programming technique to obtain the optimal segmentation for $m+1$ change points reusing the information that was calculated for the m change points
- Procedure was popularised in the work of Bai & Perron (1998, 2003)
- While this algorithm is exact, the computational complexity is considerably higher than that of binary segmentation
- The pruned exact linear time (PELT) algorithm also provides an exact solution, but is more computationally efficient

Changes in mean

- To provide an example of the output from these tests we make use of simulated data
  - Draw 360 observations from a random normal distribution with a constant variance
  - First 100 observations have a mean value of zero
  - Second 50 observations have a mean of 1.5
  - Third 90 observations have a mean of zero
  - Last 120 have a mean value of -0.8
- This data is then used to test for changes in the mean

Figure 1: Simulated time series with different means

Figure 2: Binary segmentation - change points at observation 85, 152, & 233

Figure 3: Segment neighborhood method - change points at obs 100, 149 and 233

Figure 4: PELT method - change points at observation 100, 149 and 233

Changes in variance

- To test for changes in the variance
  - Draw 360 observations from a random normal distribution with a constant mean
  - First 100 observations have a variance of 1
  - Second 50 observations have a variance of 2
  - Third 90 observations have a variance of 1
  - Last 120 observations have a variance of 0.5

Figure 5: Change in variance - observations 104, 150 and 236

Changes in mean and variance

- To test for changes in the mean and the variance we combine the changes that pertain to the first two moments of the data
- We assume that the data is governed by a normal distribution

Figure 6: Change in mean & variance - observations 107, 150 and 239

Structural break tests

- Term structural break is synonymous with change point, but would usually refer to cases where there is a change in the regression coefficients
- These results are conditional on the specification of model that is under consideration
- For example, consider the linear regression model

$$y_t = x_t^\top \beta_j + \varepsilon_t \quad j = \{1, \ldots, m+1\}$$

- The hypothesis that the regression coefficients are constant may be constructed as

$$H_0 \ : \ \beta_j = \beta_0$$

- Against the alternative that at least one coefficient varies over time
- If x_t takes the form of a constant, then the above test would be equivalent to testing for a change point in the mean

Structural break tests

- To test such a hypothesis we could also make use of information that is contained in the residual, ε_t

- Construct F-statistics to consider whether or not a change occurred at time τ use the residuals for a subsample $\hat{\varepsilon}(\tau)$ are compared to the residuals for the unsegmented dataset, $\hat{\varepsilon}$

$$F_i = \frac{\hat{u}^\top \hat{u} - \hat{u}(\tau)^\top \hat{u}(\tau)}{\hat{u}(\tau)^\top \hat{u}(\tau) / (T - 2k)}$$

- When the date of the potential structural break is unknown, one would test all possible positions for τ, where the null hypothesis is rejected if the supremum is too large

- Hansen (1997) provides critical values that are used for the approximate asymptotic p values for this test

Structural break tests

- As an alternative, a generalised fluctuation test framework seeks to identify departures from constancy in a graphical way
- Considers the behaviour of fluctuations in either the residuals or the parameter estimates
- Popular variants include the CUSUM test that considers the cumulated sums of the residuals from a particular model
- If they represent white noise then it would be expected that they should be centered on zero
- However, where these residuals display a significant departure from zero, then this behaviour may suggest that there is a structural break in the data

Chow's breakpoint test

- Chow (1960) breakpoint test seeks to fit the same regression model to separate sub-samples of the data

- Investigates whether the null hypothesis of "no structural change" holds after constructing a F-test statistic for the parameters in the two models

- To develop some intuition for this test consider the simple linear regression model

$$y_t = x_t^\top \beta_j + \varepsilon_t$$

- Would allow for the possibility that β_j may be time varying, in that it may take on two possible values

Chow's breakpoint test

- To test whether or not the coefficient estimate changes at date τ we could consider

$$\beta_j = \begin{cases} \beta & t \leq \tau \\ \beta + \delta & t > \tau \end{cases}$$

- If the break date is known then the null hypothesis of no break, $\delta = 0$, may be considered against the alternative of a nonzero break, $\delta \neq 0$
- Equivalent to testing the hypothesis that the coefficient δ is zero in the regression

$$y_t = x_t^\top \beta_t + \delta Z_t(\tau) + \varepsilon_t$$

- Test can be computed using conventional t-statistics where the hypothesis of no break is rejected at the 5% significance level if the absolute value is greater than 1.96
- Alternatively, we could test for a break at this point with the aid of an *F*-statistics
- Major drawback of this procedure is that the change point must be known *a priori* as it is required to split the sample into two sub-samples

Quandt likelihood ratio test

- Quandt (1960) likelihood ratio (QLR) test is a natural extension of the Chow test where a F-test statistic is calculated for all potential breakpoints within an interval
- The interval is usually dependent upon the degrees of freedom that are required for the estimation of the regression model
- The largest test statistic across the grid of all potential break points is then identified as the QLR statistic
- One would reject the null hypothesis of no structural change if the absolute value of this test statistic is relatively large
- Andrews (1993) and Andrews & Ploberger (1994) provide appropriate asymtotic p -values for this statistic
- This test is usually performed as a sup$F$-test and displays good power against the alternative of a breakpoint

Testing for multiple structural breaks

- Bai & Perron (1998, 2003) extend this approach to test for multiple structural breaks
- They make use of *F*-tests for 0 vs. τ_1 breaks and τ_1 vs. τ_2 breaks, etc.
- For m possible breakpoints they consider the least squares estimates for the different possible β_j coefficients to identify the coefficients that minimise the resulting RSS
- Information criteria are often used for model selection, which would identify the number of breakpoints
- Bai & Perron (2003) suggest that the AIC usually overestimate the number of breaks and as such the BIC is usually preferred
- To perform this calculation in a relatively expedient manner they make use a dynamic programming procedure

CUSUM test

- CUSUM test is based on the cumulative sum of the recursive residuals that utilises a generalised fluctuation test framework
- Makes use of a plot for the cumulative sum together with the 5% critical boundaries
- Test suggests there is parameter instability if the cumulative sum breaks either of the two boundaries

Examples of structural breaks

- To provide an example we can make use of simulated data once again
- In this case we are going to assume that the data displays a certain degree of persistence
    - First 100 observations represent an AR(1) process with a single coefficient of 0.9
    - Second 100 observations represent a MA(1) process with a coefficient of 0.1
    - Last 100 observations represent and ARMA(1,1) with coefficients of 0.5 and 0.3

Figure 7: Simulated time series with varying persistence

Figure 8: Quandt likelihood ratio test

Figure 9: CUSUM test

Figure 10: Bai & Perron (2003) structural break test

Conclusion

- Change point tests and structural break models seek to identify permanent changes that may arise in the first two moments of the data or the parameters of models
- Several widely used economic and financial indicators have a number of potential structural breaks, particularly over recent periods of time
- Failing to recognise this feature of the data can lead to invalid conclusions about the relevant features of the data and the forecasts from such models would usually be relatively inaccurate